Computer-aided Diagnosis of Pulmonary Infections Using Texture Analysis and Support Vector Machine Classification

Jianhua Yao, PhD, Andrew Dwyer, MD, Ronald M. Summers, MD, PhD, Daniel J. Mollura, MD

Rationale and Objectives: The purpose of this study was to develop and test a computer-assisted detection method for the identification and measurement of pulmonary abnormalities on chest computed tomographic (CT) imaging in cases of infection, such as novel H1N1 influenza. The method developed could be a potentially useful tool for classifying and quantifying pulmonary infectious disease on CT imaging.

Materials and Methods: Forty chest CT examinations were studied using texture analysis and support vector machine classification to differentiate normal from abnormal lung regions on CT imaging, including 10 patients with immunohistochemistry-proven infection, 10 normal controls, and 20 patients with fibrosis.

Results: Statistically significant differences in the receiver-operating characteristic curves for detecting abnormal regions in H1N1 infection were obtained between normal lung and regions of fibrosis, with significant differences in texture features of different infections. These differences enabled the quantification of abnormal lung volumes on CT imaging.

Conclusion: Texture analysis and support vector machine classification can distinguish between areas of abnormality in acute infection and areas of chronic fibrosis, differentiate lesions having consolidative and ground-glass appearances, and quantify those texture features to increase the precision of CT scoring as a potential tool for measuring disease progression and severity.

Key Words: Computer-aided diagnosis; pulmonary infection; texture analysis.

In trying to increase the clinical utility of infectious disease imaging, researchers currently face several challenges, including the relatively low specificity for diagnosing pathogens and the limited quantification of disease burden for assessing severity and predicting outcomes. The low specificity of infectious disease imaging stems from the similarity between visual appearances of infectious and inflammatory diseases (1). The second major challenge, the quantification of severity through radiologic techniques, requires standardized methods for measuring lesions and translating those measurements into validated clinical implications. A third unsolved problem is that the detection of subtle pulmonary parenchymal changes may not be visually apparent, and traditional visual scoring methods for pulmonary disease on computed tomographic (CT) imaging are often limited by interobserver biases and lack of validation. These limitations came to light during the outbreak of novel swine-origin influenza A/H1N1 in 2009 and 2010 (2). Reports indicated that severe infection with novel H1N1 demonstrate patchy ground-glass opacities with consolidations on thoracic CT scans (2–4). Unfortunately, this visual appearance is so similar to other infectious and inflammatory etiologies that it is difficult to unequivocally diagnose H1N1 on the basis of CT findings alone (2,5). Although most cases of H1N1 were predominately mild in severity, with a mortality of <1%, the severe cases often rapidly led to respiratory impairment and death, and it was clinically and radiologically challenging to prognosticate and identify these severe cases for earlier treatment (6).

In this report, we present a pilot method for detecting and quantifying H1N1 pulmonary infection using computer-assisted texture analysis and support vector machines (SVMs). To our knowledge, H1N1 pulmonary infection and associated inflammation have not been characterized using texture analysis to date. Simply defined, texture analysis quantifies an image by identifying statistical relationships among the pixels' densities, which can be used to identify lesions and quantify the volume

Acad Radiol 2011; 18:306–314

From the Center for Infectious Disease Imaging (CIDI) and Department of Radiology and Image Sciences, Clinical Center, National Institutes of Health, 10 Center Drive, Bethesda, MD 20892. Received August 11, 2010; accepted November 9, 2010. This research was supported by the Intramural Research Program of the National Institutes of Health, Clinical Center. **Address correspondence to**: D.J.M. e-mail: molluradj@mail.nih.gov

doi:10.1016/j.acra.2010.11.013

of an organ manifesting those patterns associated with lesions. It has been established that specific tissues and even specific pathologies yield unique texture patterns on chest CT images. Therefore, these textures could be important attributes for characterizing and distinguishing objects, lesions, and regions (7).

Several texture-based recognition systems have been developed to take advantage of these differences and classify pulmonary tissue patterns. Sluimer et al (8) summarized the progress through 2006 in lung segmentation, registration, nodule detection, tissue quantification, and classification for assessing emphysema, pulmonary nodules, pulmonary embolism, and airway pathology. Van Rikxoort et al (9) proposed an automatic method to segment the lungs, lobes, and pulmonary segments, an important development for measuring lung volumes and locating lesions. The lobes are segmented with high accuracy through voxel classification on the basis of texture properties and relationships to fissures. Uppaluri et al (10) introduced the adaptive multiple feature method, which assesses up to 22 texture features for the classification of pulmonary tissue patterns into six categories. Zavaletta et al (11) proposed a three-dimensional texture analysis method based on morphology. Caban et al (12) investigated the effects of texture block selection, data reduction, and texture smoothing in a computer-aided diagnosis (CAD) system for lung fibrosis. Lee et al (13) compared three classification techniques in lung CAD systems and concluded that SVMs achieved the best performance. Yao et al (7) developed a lymphangioleiomyomatosis detection and classification method on the basis of statistical texture analysis and machine learning techniques.

In this study, we developed a texture-based detection and classification method to quantitatively differentiate pulmonary lesions in patients infected with novel swine-origin influenza A/H1N1 from visually normal lung parenchyma in CT data. To further assess this method, we also compared the texture features in these patients to the CT texture patterns in individuals with normal lungs, fibrosis, and several infections other than H1N1, including mycobacterium avium complex (MAC), parainfluenza, and bacterial pneumonia. We hypothesized that texture analysis of CT data can accurately identify areas of abnormal lung on CT imaging in patients with H1N1 using features that are quantifiably different from those of normal lungs, fibrosis, and other infections.

MATERIALS AND METHODS

Data Sets

The data set included 40 chest CT studies, summarized in Table 1. The H1N1-infected cohort included four patients with autopsy-proven (by histology, immunohistochemistry, and reverse transcriptase polymerase chain reaction) infections. In each of these cases, the patient underwent chest CT imaging without contrast, performed in helical mode with 5-mm slice thickness, body filter, and lung kernel (6). Each of these patients had pathology and immunohistochemistry for proof of infection. The pulmonary fibrosis cohort was composed of 20 individuals confirmed by pathology and clinical data to have fibrosis secondary to lymphangioleiomyomatosis. To identify interstitial lung disease, thin-cut (slice thickness, 1–1.25 mm) high-resolution CT scans of the chest in the prone position were performed using a high-resolution technique with 1-mm slice thickness at 3-cm intervals. The normal CT cohort included 10 individuals from a screening study. Thin-cut (slice thickness, 1–1.25 mm) high-resolution CT scans of the chest were used. Four cases of infection other than H1N1 were included in this study to assess whether infections other than influenza present distinguishable texture patterns. The institutional review board approved our retrospective sampling and experimental design.

TABLE 1. Summary of Patient Population

Cohort	Number of Patients (Male/Female)	Age (y), Range (Mean ± SD)
H1N1	4 (3/1)	31–59 (50 ± 13)
Fibrosis	20 (11/9)	30–79 (55 ± 12)
Normal	10 (4/6)	38–75 (53 ± 12)
Infection other than H1N1		
Pneumonia	3 (3/0)	17–44 (33 ± 14)
Parainfluenza	2 (1/1)	59–60
MAC	1 (1/0)	82
All	40 (17/23)	17–82 (53 ± 14)

MAC, mycobacterium avium complex; SD, standard deviation.

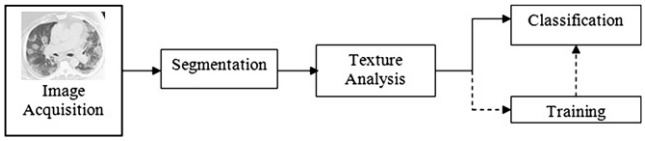

Figure 1. Method diagram for image analysis.

Method Overview

Given a chest CT data set, the proposed CAD system first segments the lungs and then analyzes the texture patterns. A classifier is then trained to distinguish regions of abnormal lung region from visually normal lung regions. Figure 1 presents a diagram of the proposed system.

Lung Segmentation

We have developed a segmentation technique that, using region growing and dynamic programming, accurately segments the left and right lungs. First, the trachea is located and segmented out, and on the basis of the trachea's position, two seed points are placed in the right and left lungs. Then, by means of a three-dimensional region-growing algorithm, the seeds are expanded to segment the entire area of the lungs. A histogram-based thresholding technique is used to refine the segmentation. Finally, a rolling-ball algorithm is applied to smooth the lung boundary. As part of the segmentation process,