

Content-based image-retrieval system in chest computed tomography for a solitary pulmonary nodule: method and preliminary experiments

Masahiro Endo · Takeshi Aramaki · Koiku Asakura · Michihisa Moriguchi · Masahiro Akimaru · Akira Osawa · Ryuji Hisanaga · Yoshiyuki Moriya · Kazuo Shimura · Hiroyoshi Furukawa · Ken Yamaguchi

Received: 16 June 2011 / Accepted: 19 December 2011 / Published online: 19 January 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract

Purpose The aim of this study was to develop a new diagnostic support system using content-based image-retrieval technology. In this article, we describe the mechanism and preliminary evaluation of this system for use with CT images of solitary pulmonary nodules.

Materials and methods With the approval of the institutional review board of Shizuoka Cancer Center, we built a database that included CT images of 461 solitary pulmonary nodules. With this database, we developed a system that automatically extracts the pulmonary nodule when the nodule area is clicked, retrieves previous cases based on an image analysis of the extracted lesion, and generates reports of the pulmonary nodule semi-automatically. We compared the percentage of correct diagnoses with and without the system using 30 solitary pulmonary nodules,

which were not included in the database, with one radiologist and two residents. As a per-user evaluation, the number of clicks required to extract the nodule region and the extracted regions was compared, and presented candidate cases were evaluated. As an evaluation of the retrieval results, the presented candidate cases were evaluated by comparing the number of diagnostic matches (benign/malignant) between the queries and four presented cases. Additionally, to evaluate the validity of the retrieval technology, the radiologist selected the most similar cases presented by the system and evaluated the visual similarity on a five-point scale.

Results With this system, the percentage of correct diagnoses for the radiologist improved from 80 to 93%. For the two residents, the diagnostic accuracy improved from 66.7 to 80% and from 76.7 to 90%, respectively. The evaluation of the number of clicks required indicated that for 19 cases with the radiologist and 12 and 11 cases with the two residents, respectively, only one click was required to extract the region. When the extracted regions were compared between the radiologist and the residents, 22 and 19 cases had a Dice's Coefficient of 0.85 or higher, respectively. For the radiologist, the number of cases that matched the diagnosis (benign/malignant) averaged 3.7 ± 0.5 among 23 malignant cases and 1.7 ± 1.4 among 7 benign cases, while for the residents, these values were 3.6 ± 0.5 and 1.1 ± 0.9 , and 3.4 ± 0.8 and 1.1 ± 1.3 , respectively. With regard to visual evaluations by the radiologist, there were 15 similar cases and 11 somewhat similar cases.

Conclusion These results suggest that, despite some differences in the search results among the users, this system has been confirmed that it can improve the accuracy of diagnosis as it displays similar cases at high probability. In addition, with the use of this system, past cases

M. Endo (✉) · T. Aramaki · K. Asakura · M. Moriguchi · M. Akimaru
Division of Diagnostic Radiology, Shizuoka Cancer Center,
1007 Shimonagakubo, Nagaizumi-cho, Sunto-gun,
Shizuoka 411-8777, Japan
e-mail: m.endo@scchr.jp

A. Osawa · R. Hisanaga · Y. Moriya · K. Shimura
Fujifilm Corporation, 2-26-30 Nishiazabu, Minato-ku,
Tokyo 106-8620, Japan
e-mail: ryuji_hisanaga@fujifilm.co.jp

H. Furukawa
Shizuoka Aoinomori Clinic,
193-1 Yanagi-cho, Aoi-ku, Shizuoka-shi,
Shizuoka 420-0007, Japan

K. Yamaguchi
Shizuoka Cancer Center,
1007 Shimonagakubo, Nagaizumi-cho,
Sunto-gun, Shizuoka 411-8777, Japan
e-mail: k.yamaguchi@scchr.jp

and their reports can be effectively referred to. Therefore, this diagnostic-assistant system has the potential to improve the efficiency of the CT image-reading workflow.

Keywords Content-based image-retrieval · Pulmonary nodule · Chest CT

Introduction

As Picture Archiving and Communication Systems (PACS) have become more popular, the work environment of radiologists has changed dramatically. In addition, over the past several years, diagnostic imaging techniques have proliferated, from plain X-rays to CT, MRI, PET, and other modalities. Radiologists are facing enormous workloads due to an exponentially increasing volume of images from various modalities. On the other hand, the shortage of radiologists has sometimes reached near-crisis levels [1]. An effective solution is needed to improve diagnostic efficiency and prevent clinical misdiagnoses. Furthermore, education and training in diagnostic imaging modalities and methods have grown increasingly diverse and complex [2], which makes it difficult to acquire the knowledge needed during limited training periods. Therefore, more effective and efficient learning environments are needed.

To counter this problem, computer-aided diagnosis and detection systems have been proposed to support diagnoses made by radiologists. Mammography-based detection systems for candidate breast cancers [3–5], CT-based detection systems for candidate lung cancers [6], and CT-based detection systems for colon polyps [7] all detect candidate lesions to improve the efficiency of lesion detection by personnel who are responsible for diagnosis and to prevent misdiagnosis by less-experienced physicians. However, the ideal goal of achieving 100% detection raises various technical difficulties, and these systems are generally used more as safeguards against misdiagnosis. Generally, they are considered to have a relatively low impact on efficiency [6–10]. While systems that retrieve similar past cases during reading [11] or that diagnose benign/malignant lesions [12] have been proposed, their main purpose is to support less-experienced physicians in making differential diagnoses. Any improvements in diagnostic efficiency among experienced radiologists are believed to be minimal.

Systems that incorporate structured reports also exist to facilitate the preparation of diagnosis reports [13]. Such systems may help make reports more consistent and improve their quality, but the physician must still manually enter all of the information [14]. All of these systems support physicians at some specific stage of the reading workflow, whether in candidate lesion detection, lesion differentiation, or report

preparation, rather than throughout the entire reading workflow.

We developed a prototype for a new system that automatically retrieves cases that are similar to a new case and generates CT reports semi-automatically, using an existing CT database of pulmonary nodules in which the contents of structured diagnosis reports are tagged [15]. For many diseases, including cancer, among different patients, as long as the extent of the disease (the disease stage, in the case of cancer) is similar, the contents of the diagnostic reports (e.g. for pulmonary CT) are also generally similar, with only slight variations in size and location. Thus, if a template for structured diagnosis reports could be prepared for a particular disease, reports could be prepared simply by providing certain information for each patient.

In contrast to earlier systems, the present system, while limited to pulmonary nodules, is designed to support the workflow of image-reading by retrieving similar past cases to aid in image-interpretation and generation of reports semi-automatically to ease report input tasks. It also effectively uses massive amounts of examination data by compiling a database of past images, reports, and verified diagnoses.

For this study, we examined diagnoses based on CT images using the present system. This paper presents and discusses the results.

Materials and methods

Case database construction

With the approval of the institutional review board (IRB) of Shizuoka Cancer Center, we built a case database, which included data (CT images, diagnosis reports, and confirmed diagnoses) from 439 patients with 461 solitary pulmonary nodules in whom a final diagnosis had been confirmed from September 2002 to February 2008. The patients included 222 men and 217 women who ranged in age from 21 to 96 years (average 68.7 years). The cases included 396 lung cancers (326 adenocarcinomas, 37 squamous cell carcinomas, 16 large cell carcinomas, 5 adenosquamous carcinomas, and 12 others; they were considered to be almost similar to the distribution of operations for lung cancer in Japan [16].) and 65 benign lesions (inflammatory nodules and granulomas). All cases of cancer and 30 benign lesions were diagnosed by surgery, and the rest of the benign lesions were confirmed by two-year follow-ups.

Nodule with internal opacity consisted of 205 solid nodules, 157 part-solid nodules, and non-solid nodules, ranging in size from 5.0 to 62.3 mm, with an average of 22.3 mm. (Table 1). CT examinations were performed with 16-detector CT scanners (Aquilion 16; Toshiba Medical Systems, Tokyo, Japan). Scans were performed from the lung apex to the

Table 1 Nodule characteristics

Nodules	Number
For the database	461
Type	
Lung cancer	396
Adenocarcinoma	326
Squamous cell carcinoma	37
Large cell carcinoma	16
Adenosquamous carcinoma	5
Others	12
Benign	65
Inflammation	65
Internal opacity	
Solid	205
Part-solid	157
Non-solid	98
Diameter	5.0–62.3 mm (mean; 22.3 mm)
For test cases	30
Type	
Lung cancer	23
Adenocarcinoma	17
Squamous cell carcinoma	4
Large cell carcinoma	1
Adenosquamous carcinoma	1
Benign	7
Inflammation	7
Internal opacity	
Solid	14
Part-Solid	10
GGO	6
Diameter	8.3–44.9 mm (mean; 21.5 mm)

diaphragm at the end of suspended inspiration. The scan parameters were 120 kV, 135–165 mAs, and 16*0.5 mm collimation. During scanning, 100 ml of contrast material (Iopamiron 300, Bayer-Schering Pharma, Osaka, Japan) was administered intravenously. All CT images were obtained with a scan delay of 90 s and reconstructed to contiguous 5-mm slice images.

System description

This system makes it possible to display similar cases by selecting them from among the pulmonary nodule data in descending order with the image-retrieval technology. In addition, it can generate text on findings of pulmonary nodules by using image analysis (2D) and comparing the results to those in cases with similar findings. These procedures are outlined in Fig. 1.

Database construction

The database for the present system consisted of three components: CT images, diagnosis reports, and confirmed diagnoses (findings of the pulmonary nodule in a structured text format, supplemental findings in a free-text format and diagnostic impressions). These data are registered in the database and linked to the feature indices as described later that are used for content-based image-retrieval.

The reports are stored in a template format. In this template, the lesion location (lung field, lobe, segment, proximity to the pleural membrane), size, shape, border, concentration, and diagnosis are selected from a list. From the template, a structured text format composed of selected items is generated. (e.g., in the (left lung) (upper lobe S2) (directly below the peripheral pleura), a (20 × 18-mm diameter) (circular) (solid) (nodule) with a (clear border) (is seen)). Supplemental findings can be entered in a free-text format as needed.

Automatic image analysis

The user first identifies a pulmonary nodule on a patient's CT image and clicks the lesion using a computer mouse. Through the use of machine-learning techniques, the system automatically extracts the border area of the pulmonary nodule [17].

If the border area does not match the pulmonary nodule, the user clicks again until the correct border is identified. This initiates the extraction of the nodule, which takes into account all the clicked points, and displays the corrected border area. This sequence can be repeated any number of times. Once the nodule lesion is confirmed, the image-analysis system automatically extracts the location of the lesion in the lung along with certain feature indices of the lesion image when the user clicks a search button. The current system can determine the lung field (right/left) of the lesion. Feature indices are calculated by the image-analysis system and represent important radiological findings of nodules (Table 2).

Radiological findings in feature indices of internal properties include ground-glass opacity (GGO), mixed GGO, solid, and presence of cavities and calcifications. To differentiate these radiological findings, the mean, standard deviation, maximum and minimum CT values on the extracted border area are used as feature indices. Furthermore, the proportions of low-density areas and each area to the entire nodule area are used as feature indices of the cavity region of internal property based on threshold processing and labeling processing.

The second moment and circularity are also used as shapes of the nodule and marginal properties (regular or irregular).

Finally, the feature indices of marginal properties include spicular and pleural indentation, and closeness to pleura. The image values obtained by a Hessian filter for the extraction of a linear structure in the marginal region of the nodule are used

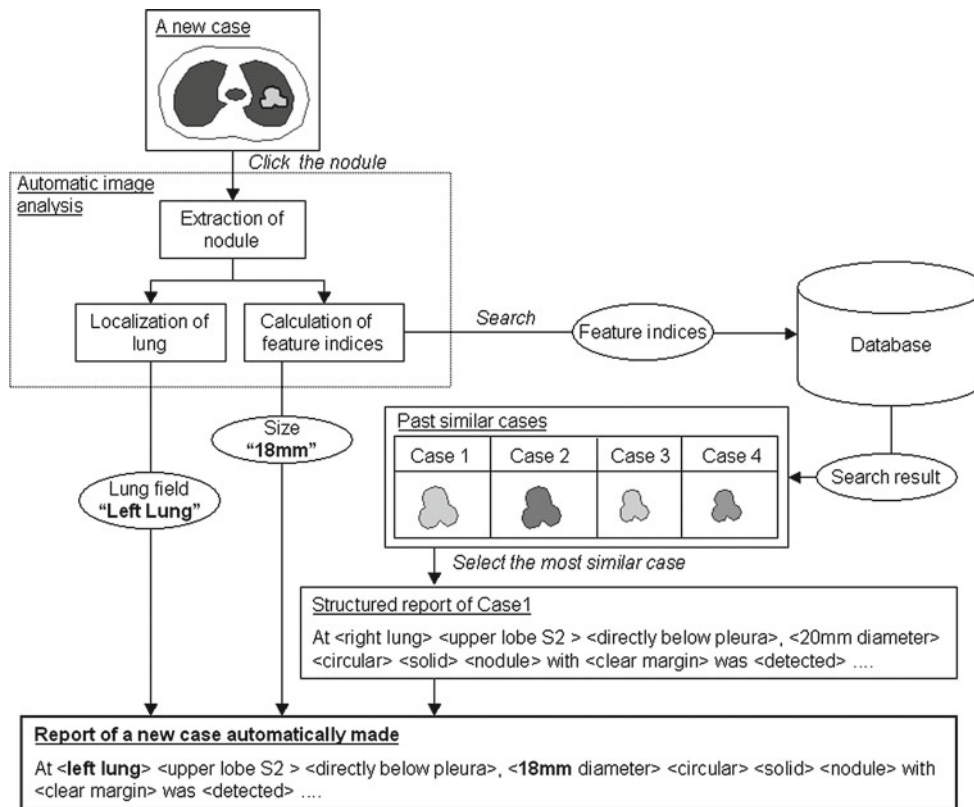


Fig. 1 Schematic diagram of the system

Table 2 Database indices

Radiological findings	Index
Shape	Area and maximum diameter
	Circularity and second moment
Internal property	Mean, SD, max and min of internal CT value
	Area and distribution of high- and low-density regions
Margin	Area and contrast of marginal fiber structures
	Area and distribution of the marginal region

to differentiate spicular and pleural indentation. The mean and standard deviation of the CT values in the marginal region of the nodule are used to determine the closeness to pleura.

Image-retrieval

The system compares the extracted feature indices for the new case and those of previous cases in the database. The similarity S , which is calculated by respective feature indices, is used for this comparison. Similarity can be expressed in terms of the Euclidean distance in the high-dimensional feature space, and the following formula (1) calculates the

similarity S_i between the newly input image and the i th registration image using n feature indices:

$$S_i = |Q - R_i| \quad Q = (q_1, q_2, \dots, q_n),$$

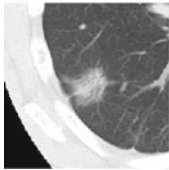
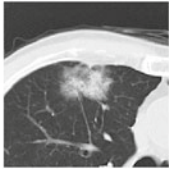
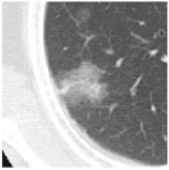
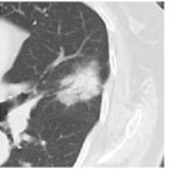
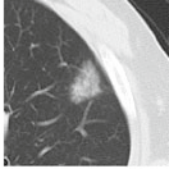
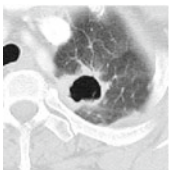
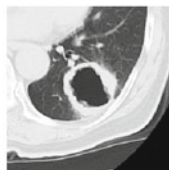
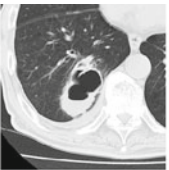
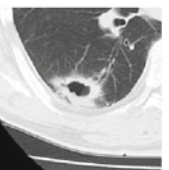
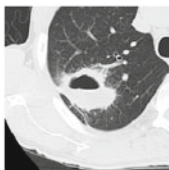
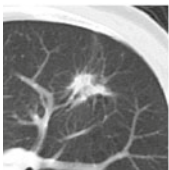
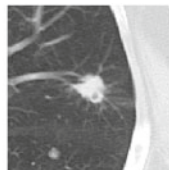
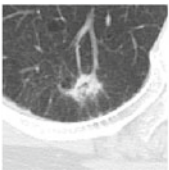
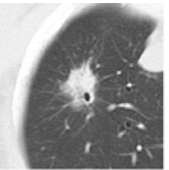
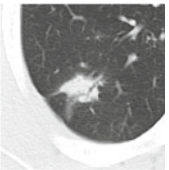
$$R_i = (r_{i1}, r_{i2}, \dots, r_{in}) \tag{1}$$

where Q is a feature vector of the input image, R_i is a feature vector of the i th registration image, and q, r are the feature mentioned in “Automatic image analysis”, which are normalized by the mean values and standard deviations of respective feature indices registered in the database.

The system then automatically retrieves the cases that are the most similar from the database and presents thumbnails of a set number of cases in decreasing order of relevance (similarity). In the present study, the number of cases retrieved by the system was set to four.

Table 3 shows examples of the retrieval results. It shows similar cases for three queries: (a) peripheral nodules with ground-glass opacity, (b) pleural masses with cavity, and (c) irregular and spicular nodules with air bronchogram. However, since it determines similarity only by image analysis, a case with a different diagnosis may also be retrieved as shown in query (c).

Table 3 Examples of retrieval results

Query	Similar cases			
	No.1	No.2	No.3	No.4
 (a) Adenocarcinoma	 Adenocarcinoma	 Adenocarcinoma	 Adenocarcinoma	 Adenocarcinoma
 (b) Squamous cell carcinoma	 Squamous cell carcinoma	 Squamous cell carcinoma	 Squamous cell carcinoma	 Squamous cell carcinoma
 (c) Adenocarcinoma	 Adenocarcinoma	 Inflammation	 Adenocarcinoma	 Adenocarcinoma

Text generation on findings

Among the four similar cases presented, the user selects the case that is the most similar to the new case, and an input template window for the report of the new case appears. At this stage, the lesion size and affected lung field are automatically entered based on the results of image analysis, and the remaining items are assigned to be those of the selected reference case. The user can modify the selected items if necessary, and text on the findings, consisting of structured items, is generated when the user presses a text-making button. The user can enter any supplementary findings and diagnostic impressions in a free-text format on an as-needed basis.

Database registration

The database registration follows the same steps as in the search operation. First, as well as the search method described above, the user clicks on a lesion on a CT image and makes the necessary corrections to finalize the border area of the lesion. Once the nodule area is confirmed (by clicking the registration button), the image-analysis system automatically extracts feature indices and the affected lung

field (right or left). After the lesion area is confirmed, an input template window is used to prepare the structured report of the lesion appears. The lesion size and affected lung field are also entered automatically, and the user enters items in the structured report. The registration is finally completed after supplementary findings, if needed, are entered in a free-text format. Pressing the input complete button on the template window registers the image, the border area of the lesion, feature indices, and text of findings.

System evaluation

Reading experiments were performed to clinically evaluate diagnostic accuracy. The observers included one radiologist (initials: ME) and two residents (respiratory disease specialists learning lung cancer; resident A: AT, 7-year career and resident B: SI, 10-year career). The test cases consisted of 30 cases with confirmed diagnoses that were not included in the database (17 adenocarcinomas, 4 squamous cell carcinomas, 1 large cell carcinoma, 1 adenosquamous carcinoma, and 7 inflammations). CT images of the test cases were taken under the same conditions as those registered in the database. All cases were diagnosed after surgery.

Table 4 Results regarding diagnostic accuracy (benign and malignant)

Observer	Percentage of correct answers					
	Malignant (<i>n</i> = 23)		Benign (<i>n</i> = 7)		Total (<i>n</i> = 30)	
	Without CAD	With CAD	Without CAD	With CAD	Without CAD	With CAD
Radiologist	73.9% (17/23)	91.3% (21/23)	100% (7/7)	100% (7/7)	80% (24/30)	93.3% (28/30)
Resident A	56.5% (13/23)	78.3% (18/23)	100% (7/7)	85.7% (6/7)	66.7% (20/30)	80% (24/30)
Resident B	73.9% (17/24)	100% (23/23)	85.7% (6/7)	57.1% (4/7)	76.7% (23/30)	90.0% (27/30)

In the reading experiments, each observer read the CT images of the test cases and prepared diagnosis reports with or without the aid of the present system. To evaluate diagnostic accuracy, the confirmed diagnoses (benign/malignant) of the test cases were compared to those determined with and without the present system. Also, to compare diagnoses of benign/malignancy, the percentage of correct diagnosis was calculated for all 30 test cases and separately for the 23 malignant and 7 benign cases for each of the three observers. The experiments with or without the present system were performed at an interval of at least one month. The order of presentation for the 30 test cases was randomized.

As a per-user evaluation, the number of clicks required for marking and the marked regions was compared and presented candidate cases were evaluated. The marked regions were compared between those marked by the residents A and B and those by the radiologist. As an indicator of the comparison, Dice's coefficient that measures the degree of similarity between two sets was used. Dice's coefficient (*R*) between each nodule region is calculated as follows:

$$R = \frac{2|X \cap Y|}{|X| + |Y|} \quad (2)$$

where *X* and *Y* are the regions marked by the radiologist and the resident, respectively [18].

As an evaluation of the retrieval results, each observer selected the case that was the most similar to the query from among the four retrieved similar cases (Nos 1–4) for comparison among each observer. Moreover, the presented candidate cases were evaluated by comparing the number of diagnostic matches (benign/malignant) between the query as against the four presented cases.

Additionally, to evaluate the validity for the similar previous cases retrieved by the present system, after the above experiments were completed, the radiologist selected the case that was most similar to the test case from among the four cases chosen by the system and visually evaluated the level of similarity by assigning one of five grades (5: similar, 4: somewhat similar, 3: neither similar nor dissimilar, 2: somewhat dissimilar, and 1: dissimilar).

Table 5 Number of clicks

Clicks	Radiologist	Resident A	Resident B
1	19	12	11
2–5	5	8	2
6–9	0	3	10
10–	6	7	7

Results

Table 3 shows the results of a comparison of diagnostic accuracy as part of the verification of clinical effects. The percentage of a correct benign/malignant diagnosis for the radiologist was 80% (24/30) and 93.3% (28/30) without and with the system, respectively. These values for resident A were 66.7% (20/30) and 80% (24/30), and those for resident B were 76.7% (23/30) and 90.0% (27/30). Thus, for all three observers, the diagnostic accuracy was improved with the system. For the radiologist, the accuracy of diagnosis for benign lesions was 100% (7/7) with or without the system, while the accuracy of diagnosis for malignant lesions was 73.9% (17/23) without the system and 91.3% (21/23) with the system. For resident A, the accuracy of diagnosis for benign lesions was 100% (7/7) without the system and 85.7% (6/7) with the system, while the accuracy of diagnosis for malignant lesions was 56.5% (13/23) without the system and 78.3% (18/23) with the system. For resident B, the accuracy of diagnosis for benign lesions was 85.7% (6/7) without the system and 57.1% (4/7) with the system; the accuracy of diagnosis for malignant lesions was 73.9% (17/24) without the system and 100% (23/23) with the system (Table 4).

The distribution of clicks required by each user is shown in the table. Regarding cases read by the radiologist, 19 were read without correction, 5 required 2–4 clicks, none required 6–8 clicks, and 6 required 10 or more clicks. Regarding cases read by resident A, 12 were read without correction, 8 required 2–4 clicks, 3 required 6–8 clicks, and 7 required 10 or more clicks. Regarding cases read by resident B, 11 were read without correction, 2 required 2–4 clicks, 10 required 6–8 clicks, and 7 required 10 or more clicks (Table 5).

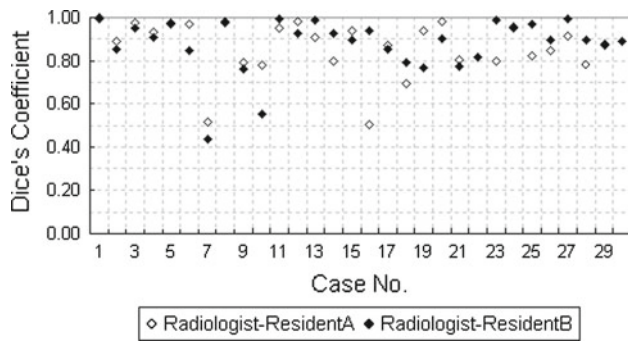


Fig. 2 Result regarding Dice’s coefficients

Table 6 The user-selected most similar case from among four retrieved cases

Observer	No. 1	No. 2	No. 3	No. 4
Radiologist	10	9	7	4
Resident A	7	4	7	12
Resident B	7	9	6	8

Figure 2 shows Dice’s Coefficients for resident A and resident B in comparison with the radiologist. As shown, 19 of resident A’s cases and 22 of resident B’s cases had a Dice’s Coefficient of 0.85 or higher. Regarding cases retrieved by the system, the average number of the same cases for the resident A and those for the radiologist was 2.47 ± 1.43 , and the average number of matched cases for resident B and those for the radiologist was 2.09 ± 1.34 .

The number of cases with a Dice’s Coefficient of 0.85 or lower was 11 for resident A and 8 for resident B. Regarding cases retrieved by the system, the average number of matched cases for resident A and those for the radiologist was 0.09 ± 0.13 and the average number of matched cases for resident B and those for the radiologist was 0.13 ± 0.35 .

Table 6 shows the results of a survey in which each observer selected the most similar case to the query among the four presented similar cases (Nos. 1–4). The number of cases selected by the radiologist as being the most similar to the query was 10 for No.1, 9 for No. 2, 7 for No. 3, and 4 for No. 4. Likewise, resident A selected 7 cases for No. 1, 4 for No. 2, 7 for No. 3, and 12 for No. 4, while resident B selected 7 cases for No. 1, 9 for No. 2, 6 for No. 3, and 8 for No. 4. Regarding the results of the radiologist, the number of cases that matched the diagnoses (benign/malignant) averaged of 3.7 ± 0.5 among 23 malignant cases and 1.7 ± 1.4 among 7 benign cases. For resident A, an average of 3.6 ± 0.5 among 23 malignant cases and 1.1 ± 0.9 among 7 benign cases matched the diagnosis (benign/malignant). For resident B, an average of 3.4 ± 0.8 among 23 malignant cases and 1.1 ± 1.3 among 7 benign cases matched the diagnosis (benign/malignant) (Table 7). Additionally, according to the

Table 7 The Average number of cases that matched the actual diagnosis

Observer	Malignant (<i>n</i> = 23)	Benign (<i>n</i> = 7)
Radiologist	3.7 ± 0.5	1.7 ± 1.4
Resident A	3.6 ± 0.5	1.1 ± 0.9
Resident B	3.4 ± 0.8	1.1 ± 1.3

radiologist, there were 15 similar cases, 11 somewhat similar cases, 2 somewhat dissimilar cases, and 2 dissimilar cases. The percentage of cases that were judged to be visually similar by the radiologist was 87% (26/30).

Discussion

In this experiment on diagnostic interpretation, we evaluated both the accuracy achieved by radiologists and residents. As a result, the percentage of questions that were answered correctly on the differential diagnoses for pulmonary nodules by the radiologist improved. This suggests that the present system could improve the diagnostic accuracy. For the residents, the percentage of correct differential diagnoses significantly improved, and in fact exceeded that for the radiologist without access to the present system, which suggests that the system may also assist residents in making accurate diagnoses. In this experiment, the number of cases that required 10 or more clicks for border extraction was 10 for the radiologist, 6 for resident A, and 7 for resident B. These are cases that required fine adjustment such as at the border of a GGO region. To reduce the time spent on reading such cases, it is also important to enhance the extraction accuracy of fine regions so that fewer clicks are required.

With the present support system, the radiologist visually evaluated the existence of a similar case in 87% of the test cases. Cases were often judged to be dissimilar (low concordance) due to the relatively few comparable registered cases in the current database, such as cases in which patients presented with organized pneumonia or inflammation in the process of healing. Similar past cases were less likely to be registered in the database for such patients, and the cases presented by the system may have included dissimilar cases that may not have been presented in a more complete database, which would require users to make their own diagnoses. Although three or more cases out of four were presented as matched diagnoses (benign/malignant) in the retrieval of malignant cases, only an average of one case was presented in the retrieval of benign cases. Under such circumstances, a less-experienced user could diagnose a benign case as malignant because the system presents malignant cases, even though the user initially diagnosed it as benign without the aid of the system. In the reading experiments described here, this was likely the reason for the relatively low percentage

of correct diagnoses for benign cases among residents. The search accuracy with the current database can be improved by adding more instances of lung cancers other than adenocarcinoma and cases of benign lesions.

Finally, the number of similar cases presented by the present system was experimentally set to four. Although the four presented cases (Nos. 1–4) were almost evenly selected by the users as the most similar case and no bias was noted, the fifth or subsequent similar cases may include the most similar case selected by other users, considering that the retrieval results varied from user to user. Consequently, an increase in the number of presented cases may increase the likelihood that the system presents similar cases. One finding that pointed to potential limitations of the current system was that sometimes, for a given case, the system presented different similar cases from user to user. This was due to the effect on image-retrieval of differences in the confirmed pulmonary nodule areas or key slices for lesion selection. Therefore, we will need to improve the lesion images used for queries, perhaps by expanding searches to include 3D data. The current study used only CT images of lung fields with a slice thickness of 5 mm and contrasted CT. Rather than including only cases with the same imaging conditions and slice thickness, we would need to include cases that included different imaging conditions and slice thicknesses. Finally, these experiments involved only three observers. Additional studies involving more observers are needed to evaluate its clinical efficacy, such as with regard to reporting time and the quality of the report.

Future studies should address not only the issues mentioned above but also other diseases so that the system can be improved to become more useful.

Conclusion

We have developed a prototype of a content-based image-retrieval system for CT images of solitary pulmonary nodules and conducted an initial evaluation. With this system, the diagnostic accuracy for a radiologist improved from 80 to 93%. For the two residents, the diagnostic accuracy improved from 66.7 to 80% and from 76.7 to 90%, respectively. Therefore, this system may be useful for improving work efficiency among experienced radiologists, and diagnostic accuracy among residents.

Acknowledgements The authors thank Drs. Satoshi Igawa and Akihiro Tamiya at the Division of Thoracic Oncology, Shizuoka Cancer Center for their assistance with the reading experiments. The present study was conducted as part of the Fuji Pharma Valley Initiatives, a business-academia collaborative project in Shizuoka Prefecture (<http://www.fuji-pvc.jp/>).

Conflict of interest There is no actual or potential conflict of interest in relation to this article.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Sethi SK (2008) Radiology in today's flat world. *Internet J Radiol* 8(1)
- William RH (2006) An opportunity for radiology. *Radiology* 238(2):389–394
- Giger ML, Huo Z, Vyborny CJ et al (2002) Intelligent CAD workstation for breast imaging using similar to known lesions and multiple visual prompt aids. *Med Imaging* 4684:768–773
- Takeo H, Shimura K, Imamura K et al (2005) Detection system of clustered microcalcifications on CR mammogram. *IEICE Trans Inf Syst* E88-D:2591–2602
- Baydush AH, Catarious DM, Abbey CK et al (2003) Computer-aided detection of masses in mammography using subregion Hotelling observers. *Med Phys* 30(7):1781–1787
- Zhanyu G, Berkman S, Chan HP et al (2003) Computer-aided detection of lung nodules: false positive reduction using a 3D gradient field method and 3D ellipsoid fitting. *Med Phys* 30(7):1592–1601
- Taylor SA, Greenhaigh R, Ilangovan R et al (2008) CT colonography and computer-aided detection: effect of false-positive results on reader specificity and reading efficiency in a low-prevalence screening population. *Radiology* 247(1):133–140
- Näppai J, Yoshida H (2003) Feature-guide analysis for reduction of false positives in CAD of polyps for computed tomographic colonography. *Med Phys* 30(7):1592–1601
- Baker JA, Lo JY, David M et al (2004) Computer-aided detection in screening mammography: variability in cues. *Radiology* 223(2):411–417
- Li F, Engelmann R, Melz CE et al (2008) Lung cancers missed on chest radiographs: results obtained with a commercial computer-aided detection program. *Radiology* 246(1):273–280
- Aisen AM, Broderick LS, Winer-Muram H et al (2003) Automated storage and retrieval of thin-section CT images to assist diagnosis: system description and preliminary assessment. *Radiology* 228:265–270
- Li Q, Li F, Shiraishi J et al (2003) Investigation of new psychophysical measures for evaluation of similar images on thoracic computed tomography for distinction between benign and malignant nodules. *Med Phys* 30(10):2584–2592
- Strucrad Web site. <http://portal.structurad.com/>. Accessed August 24 (2006)
- Langlotz CP (2000) Structured reporting in radiology. *Newsletter of Society for Health Service Research in Radiology*; Winter
- Endo M, Aramaki T, Asakura K (2009) A semi-automated chest CT diagnosis assistant using content-based retrieval of CT images with structured report. *Int J Comput Assist Radiol Surg* 4(1):S366
- Miyoshi S, Kadokura M, Kondo H, Saito Y, Haniuda M, Fujii Y (2011) 2008 nendo kokyu-ki geka shujutsu toukei (in Japanese). *J Jpn Assoc Chest Surg* 25(1):124–132
- Li Y, Hara S, Ito W et al (2007) A machine learning approach for interactive lesion segmentation. *Proc SPIE* 6512:651246–651248
- Dice LR (1945) Measures of the amount of ecologic association between species. *Ecology* 26(3):297–302