Ultrasound Computer-Aided Diagnosis (CAD) Based on the Thyroid Imaging Reporting and Data System (TI-RADS) to Distinguish Benign from Malignant Thyroid Nodules and the Diagnostic Performance of Radiologists with Different Diagnostic Experience

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

ABCDEFG 1,2,3 Zhuang Jin*
ABCDEFG 1,4 Yaqiong Zhu*
BC 5 Shijie Zhang
AE 1 Fang Xie
EF 1 Mingbo Zhang
BC 4 Ying Zhang
BC 4 Xiaoqi Tian
AG 5 Jue Zhang
AG 1,3 Yukun Luo
E 2 Junying Cao

1 Department of Ultrasound, Chinese People's Liberation Army (PLA) General Hospital, Beijing, P.R. China
2 Department of Ultrasound, General Hospital of Northern Theater Command, Shenyang, Liaoning, P.R .China
3 Medical School of Chinese People's Liberation Army (PLA), Beijing, P.R. China
4 Nankai University, Tianjin, P.R. China
5 Peking University, Beijing, P.R. China

* Zhuang Jin and Yaqiong Zhu contributed equally to this study
Corresponding Authors: Yukun Luo, e-mail: lyk_ultra@163.com, Jue Zhang, e-mail: zj_puvip@163.com
Source of support: This study was supported by the National Natural Sciences Foundation of China (No. 8177834), the Capital Clinical Characteristic Applied Research, China (No. Z18110700170000) and the Nature Science Foundation of Beijing, China (No. 7194318)

Background: The diagnosis of thyroid cancer and distinguishing benign from malignant thyroid nodules by junior radiologists can be challenging. This study aimed to develop a computer-aided diagnosis (CAD) system based on the Thyroid Imaging Reporting and Data System (TI-RADS) to distinguish benign from malignant thyroid nodules by analyzing ultrasound images to improve the diagnostic performance of junior radiologists.

Material/Methods: A modified TI-RADS based on a convolutional neural network (CNN) was used to develop the CAD system. This retrospective study reviewed 789 thyroid nodules from 695 patients and included radiologists with different diagnostic experience. Five study groups included the CAD group, the junior radiologist group, the intermediate-level radiologist group, the senior radiologist group, and the group in which the junior radiologist used the CAD system. The ultrasound findings were reviewed and compared with the histopathology diagnosis.

Results: The CAD system for the diagnosis of thyroid cancer showed an accuracy of 80.35%, a sensitivity of 80.64%, a specificity of 80.13%, a positive predictive value (PPV) of 76.02%, a negative predictive value (NPV) of 84.12%, and an area under the receiver operating characteristic (ROC) curve (AUC) of 0.87. The accuracy of the junior radiologists in diagnosing thyroid cancer using CAD was similar to that of intermediate-level radiologists (79.21% *vs.* 77.57%; P=0.427).

Conclusions: The use of ultrasound CAD based on the TI-RADS showed potential for distinguishing between benign and malignant thyroid nodules and improved the diagnostic performance of junior radiologists.

MeSH Keywords: Diagnosis • Image Processing, Computer-Assisted • Thyroid Neoplasms • Ultrasonography

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/918452

📄 3631   ▦ 5   📊 7   📚 19

Background

Thyroid nodules are a common occurrence, particularly in women. Depending on the method of examination, thyroid nodule has an incidence of 19–68% [1], and approximately 10% of patients presenting with a thyroid nodule are at risk of malignancy [2]. The incidence of thyroid cancer is 8.28 per million and continues to increase by 5% each year [3]. Among the differentiated types of thyroid carcinoma, papillary thyroid carcinoma is the most frequent histological type. Diagnostic ultrasound examination of the thyroid has the advantages of convenience and effectiveness and has been recommended by the American Thyroid Association (ATA) as the primary method for early detection and diagnosis of the thyroid nodule [1].

There are many similarities in the ultrasound features of both benign and malignant thyroid nodules. The identification of these features mainly depends on the experience of radiologists to prevent misdiagnosis, which may be more common in inexperienced or junior radiologists. Also, ultrasound has inherent limitations of interobserver and intra-observer variability, which lead to differences between operators in both image acquisition and interpretation. It is important to conduct an objective evaluation and analysis of the echogenic structure of thyroid nodules to reduce misdiagnosis [4], standardize reporting [5], identify malignant tumors [6], and avoid unnecessary biopsies [6–8]. For these reasons, the American College of Radiology (ACR) proposed an ultrasound-based risk stratification system to identify nodules that warrant biopsy or ultrasound follow-up, the Thyroid Imaging Reporting and Data System (TI-RADS) [9]. However, the classification system based on the TI-RADS is relatively complicated to implement, especially for inexperienced radiologists, and these challenges may allow misinterpretation of thyroid nodules on ultrasound to persist.

Computer-aided diagnosis (CAD) is a newly developed technique for the diagnosis of thyroid nodules based on ultrasound features. To overcome the complexity of applying ultrasound descriptors and interobserver variability in the diagnosis of thyroid nodules, recent studies have applied the CAD system to thyroid ultrasound to assist in the detection of nodules or in making decisions in clinical practice [10,11]. Studies have also shown that thyroid cancer can be diagnosed with a high degree of accuracy using the CAD system [12–15]. Gao et al. diagnosed benign and malignant nodules using CAD and assessed the findings from experienced radiologists [16]. The findings in this previous study showed that the accuracy, sensitivity, and specificity of the CAD system were 82.2%, 96.7%, and 48.5%, respectively [16]. Currently, few studies have been conducted to compare the use of the CAD system by radiologists with different levels of experience in the diagnosis of thyroid cancer. Also, improvements in diagnosis made by junior radiologists using the CAD system remain to be investigated.

Therefore, this study aimed to develop a CAD system based on the TI-RADS to distinguish benign from malignant thyroid nodules on ultrasound to improve the diagnostic performance of junior radiologists.

Material and Methods

Study design

Institutional Review Board (IRB) approval for this retrospective study was obtained from the Ethics Committee of the PLA General Hospital, and informed consent was obtained from each patient who participated in the study. All procedures were performed in accordance with the guidelines of the Helsinki Declaration.

An initial retrospective review identified 1,181 cases of thyroid nodules from 953 consecutive patients who underwent thyroid ultrasonography in the PLA General Hospital between May 2016 to June 2017. The study inclusion criteria were patients who completed preoperative ultrasound examination of thyroid nodules, and who underwent surgery or core needle biopsy (CNB) after thyroid examination, or fine-needle aspiration biopsy (FNAB) on at least two occasions within a one-year interval for benign thyroid benign lesions (except adenomas). Patients were included who underwent initial FNAB and ultrasound follow-up at >12 months after FNAB for benign thyroid lesions (except adenomas).

The study exclusion criteria included inflammatory lesions or suspicious inflammatory lesions on biopsy, or biopsy results that were unclear with no surgical resection, or patients who had poor quality images. There were 219 cases that were excluded based in the histology of FNAC diagnosis, of which, 128 cases were inflammatory lesions, and pathological findings of 91 cases were non-specific. There were 173 cases with poor-quality images.

Finally, 789 thyroid lesions in 695 patients were included in this study (Figure 1). A radiologist with 15 years of experience (author FX) selected the appropriate images for use in this study. Strict quality controls were applied for all image acquisition and histological analysis in each case.

The computer-aided diagnosis (CAD) system

Peking University developed the computer-aided diagnosis (CAD) system used in this study. The CAD system was based on the Thyroid Imaging Reporting and Data System (TI-RADS) automatic scoring system from the American College of Radiology (ACR) [9], and was based on a convolutional neural network (CNN). A multi-task learning network implemented
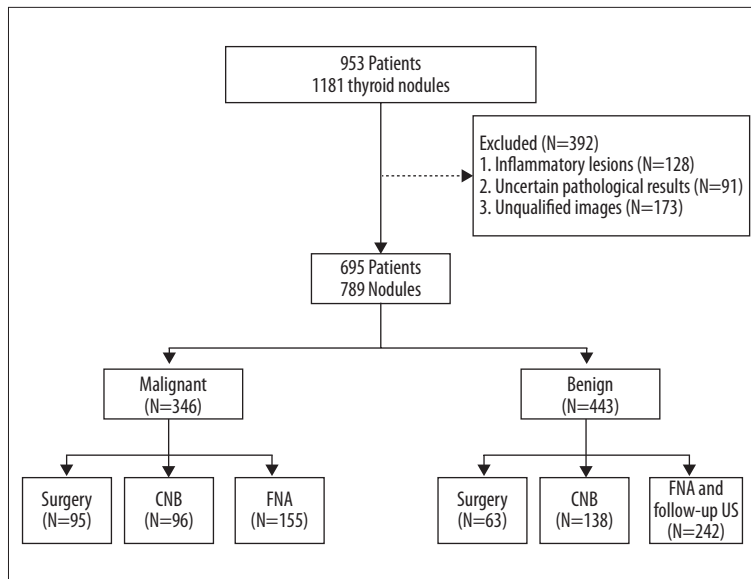
Figure 1. Flowchart of the study protocol.

the network training, and flow charts were used to explain the principle of the system (Figure 2). The radiologists manually delineated the boundary of the thyroid nodule in each image. An image preprocessing algorithm was then used to extract the input data of the network. In the image preprocessing algorithm, each image was cropped by a minimum surrounding square of the region of interest (ROI) of the nodules.

In this study, based on the convolutional neural network (CNN) architectures, a multi-task training CNN was designed to diagnose thyroid nodules, as the core of the automatic analysis. The multi-task network was used to detect clinical features. Six types of descriptors were selected as clinical features for the thyroid nodules, which were cystic, solid, mixed cystic and solid, irregular, with macrocalcifications, and with punctate echogenic foci.

Therefore, the equation for the system output included P_{tirads}=[$P1$, $P2$, …, $P6$] that denoted the predicted probabilities for the six types of clinical features. The predicted probability of each clinical feature was defined as:

$$P_n = \frac{e^{q_1}}{\sum_i^2 e^{q_i}}$$

Where q is a pair of network outputs of the clinical feature n. Therefore, the loss function of the network link was calculated as:

$$L = \sum_{n=1}^{6} Y_n log P_n + (1 - Y_n) \log (1 - P_n),$$

Where Y_n was the actual measurement (ground truth) of the dataset. Also, there were another two indices considered in the automatic analysis system, the aspect ratio R_{aspect} and echogenicity ratio $R_{echogenicity}$, which were calculated for each image:

$$R_{aspect} = \frac{w_{roi}}{h_{roi}}$$

$$R_{echogenicity} = \frac{echogenicity_{nodule}}{echogenicity_{thyroid}}$$

where w_{roi} was the width of ROI, and h_{roi} was the height of ROI. The echogenicity$_{nodule}$ was the average intensity in the ROI. The echogenicity$_{thyroid}$ was the average intensity around the ROI.

The network training was performed with stochastic gradient descent (SGD) running on NVIDIA GTX 1080Ti computer graphics. Data augmentation and dropout were applied for each image during training to avoid over-fitting.

Thyroid ultrasonography and retrospective evaluation

Thyroid ultrasound was performed with a Philips iU-22 system (Philips Health Care, Andover, MA, USA), and IU Elite (Philips Health Care, Andover, MA, USA), a Siemens ACUSON S2000 (Siemens AG, Erlangen, Germany), an ACUSON SEQUOIA 512 (Siemens Medical Solutions, Mountain View, CA, USA), a GE Vivid E9 (General Electric Healthcare, Milwaukee, WI, USA), an Esaote MyLab Twice (Esaote Group, Italy), a Hitachi HI VISION Ascendus (Hitachi Ltd, Tokyo, Japan), an Aixplorer (SuperSonic Imaging, Aix-en-Provence, France), a Mindray Resona 7 (Mindray, Shenzhen, China), and a Fenoxaparin Vinno 70 (Vinno, Suzhou, China), with a linear-array transducer (4.5–13 MHz).

Five different methods of ultrasound image analysis were performed. Multiple experienced radiologists reviewed ultrasound
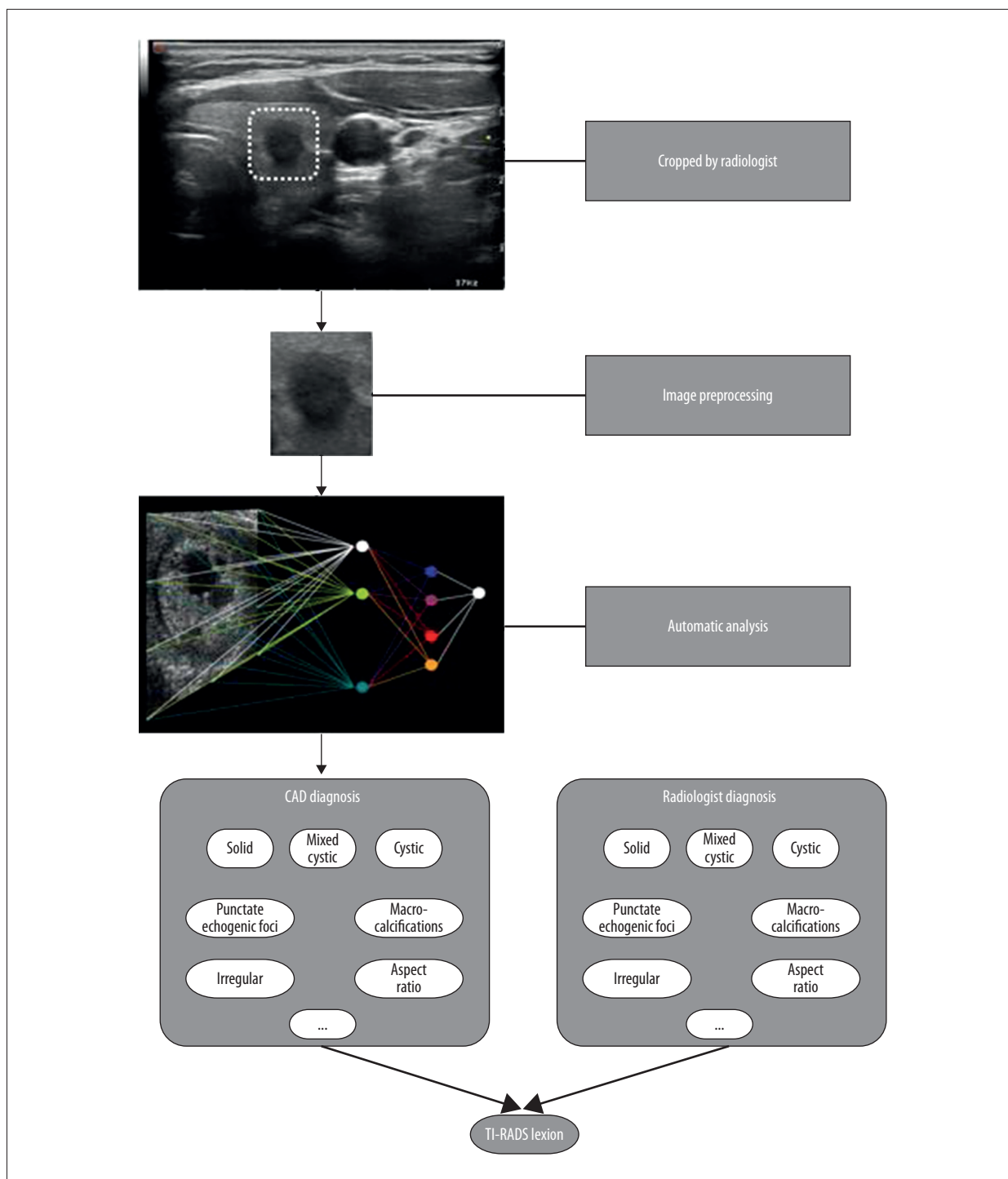
Figure 2. Flowchart of the proposed computer-aided diagnosis (CAD) system for thyroid nodule classification.

images of the thyroid nodules. One junior radiologist (SPL), with two years of diagnostic radiology experience, one intermediate-level radiologist (author, YZ), with seven years of diagnostic radiology experience, and one senior radiologist (author, MBZ), with 13 years of diagnostic radiology experience. There were five study groups that included group 1 (the CAD group),

group 2 (the junior radiologist), group 3 (the intermediate-level radiologist group), group 4 (the senior radiologist group), and group 5 (the junior radiologist and CAD group). Before image analysis, all reviewers were unaware of the pathology results and patient clinical data, and each the imaging findings.
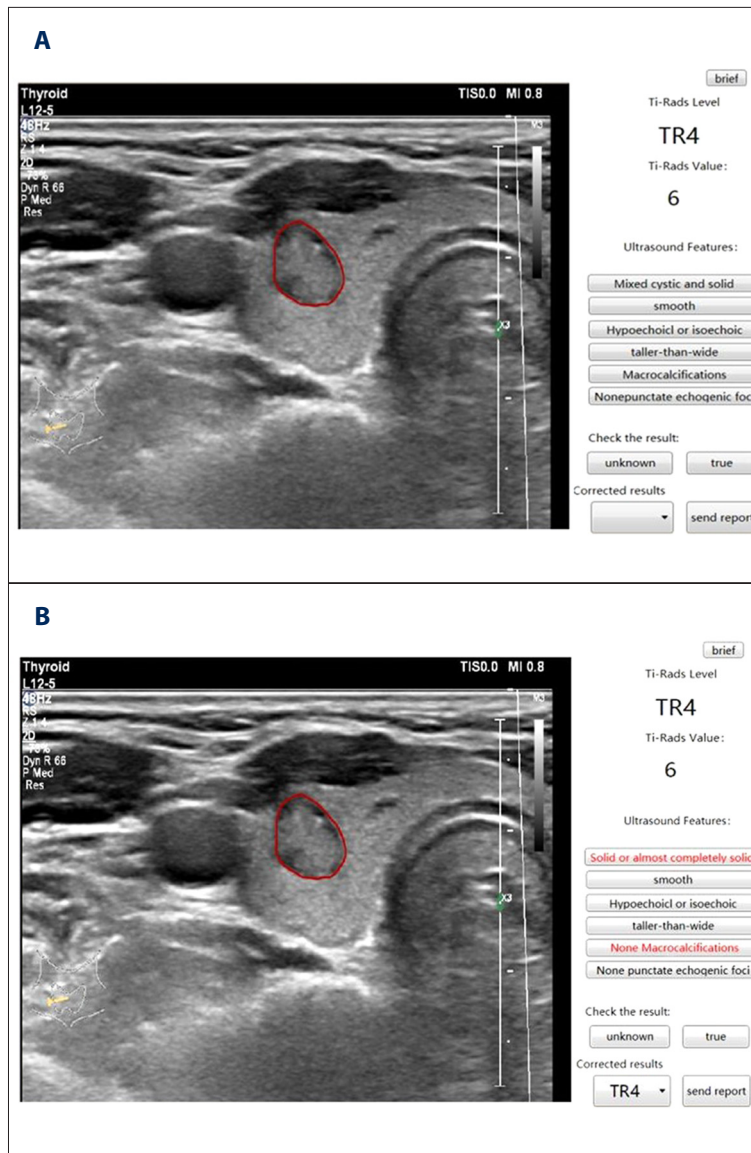
Figure 3. The ultrasound features of the Thyroid Imaging Reporting and Data System (TI-RADS) category of the thyroid nodules acquired using the computer-aided diagnosis (CAD) system. (**A**) The ultrasound features and the Thyroid Imaging Reporting and Data System (TI-RADS) category of a thyroid nodule acquired with the computer-aided diagnosis system (CAD). (**B**) The ultrasound features and TI-RADS Category of a thyroid nodule acquired in the group with the junior radiologist combined with the CAD system. The red font indicates the corrected ultrasound features of the thyroid nodule after the clinician and CAD combination. The triangle, plus sign, and dot on each receiver operating characteristic (ROC) curve indicate the performance of the junior radiologist, intermediate-level radiologist, and the senior radiologist, respectively. CAD – computer-aided diagnosis system; TI-RADS – Thyroid Imaging Reporting and Data System.

The new TI-RADS patterns published by the American College of Radiology (ACR) [9,17], thyroid nodules have the five ultrasound features of composition, echogenicity, shape, margin, and echogenic foci. The ACR TI-RADS levels were classified as TR1 (0 points, benign), TR2 (2 points, benign and not suspicious), TR3 (3 points, mildly suspicious for malignancy), TR4 (4–6 points, moderately suspicious for malignancy), and TR5 (7 points or more, highly suspicious for malignancy).

Before the start of the study, a consensus on the use of CAD systems was built by discussing their basic methods. Two separate ultrasound image review steps were performed. The first step was for a two-dimensional image review of thyroid ultrasound by group1–4. The second step, which was carried out two weeks after the first step, was to use group 5 to review the two-dimensional image of the thyroid ultrasound again.

The ultrasound features of each thyroid nodule were analyzed, based on the ACR TI-RADS lexicon and final assessment categories [17].

The boundary of the thyroid nodule in group 1 was delineated on the ultrasound image by another radiologist (author, ZJ) with 10 years of diagnostic radiology experience, and the thyroid nodule was evaluated by the CAD system. According to the TI-RADS lexicon, the CAD system automatically reported the ultrasound features of thyroid nodules, their TI-RADS grades, and the likely benign or malignant diagnosis. In group 5, the junior radiologist referenced the ultrasound features of thyroid nodules given by the CAD system to judge the benign and malignant nodules and reached a final conclusion (Figure 3).

Table 1. The clinical and demographic features of the patients in the study.

	Pathological findings		P-value
	Benign	Malignant	
Age, years (mean ± SD)	45.5±13.1	42.5±11.1	0.11
Gender			0.30
Male	90 (20.3%)	81 (23.4%)	
Female	353 (79.7%)	265 (76.6%)	
No. of nodules	443	346	
Size			<0.01
<1 cm	197	159	
1–2 cm	197	109	
>2 cm	49	78	

Statistical analysis

Data were analyzed using SPSS version 19.0 software (IBM, Chicago, IL, USA). The qualitative data were expressed as frequencies. The quantitative data were expressed as the mean±standard deviation (SD). The categorical variables were analyzed by the chi-squared (χ^2) test with Yates correction and Fisher's exact test. The accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and the area under the curve (AUC) were calculated by comparing the pathological findings. A receiver operating characteristic (ROC) curve analysis was used to compare the results of CAD system and the junior radiologists combined with the CAD system to determine the optimal cutoff value. A P-value <0.05 was considered statistically significant.

Results

Characteristics of the study population

In this study, 789 thyroid nodules from 695 patients, with an age range of 17–78 years, and a mean age of 45.3 years, met the inclusion criteria. There were 346 cases (43.85%) that were malignant, including 311 cases of papillary carcinoma, 30 cases of follicular carcinoma, and five cases of medullary carcinoma. There were 443 cases (56.15%) that were benign, including 72 cases of follicular adenoma, and 371 cases of nodular hyperplasia. Of the 789 cases of thyroid nodules, 158 cases (95 malignant, 63 benign) were confirmed by surgery, 234 cases (96 malignant, 138 benign) were confirmed by core needle biopsy (CNB), 397 cases (155 malignant, 242 benign) were confirmed by fine-needle aspiration biopsy (FNAB) (Figure 1). Thyroid nodules were divided into three groups on the basis of size. Small nodules were <1 cm, medium nodules were 1–2 cm, and large nodules were ≥2 cm. The main characteristics of the findings in the study participants are summarized in Table 1.

Diagnosis of thyroid cancer by the computer-aided diagnosis (CAD) system and by the radiologists in the different groups

Compared with the definitive histopathology diagnosis, the use of the computer-aided diagnosis (CAD) system for the diagnosis of thyroid cancer showed an accuracy of 80.35%, a sensitivity of 80.64%, a specificity of 80.13%, a positive predictive value (PPV) of 76.02%, a negative predictive value (NPV) of 84.12%, and an area under the receiver operating characteristic (ROC) curve (AUC) of 0.87. For the junior radiologist combined with the CAD system, the accuracy, sensitivity, specificity, PPV, NPV, and AUC for thyroid cancer diagnosis were 79.21%, 78.08%, 80.00%, 73.12%, 83.97%, and 0.83 respectively (Figures 4–6). The ROC curve of the Thyroid Imaging Reporting and Data System (TI-RADS) classification from the American College of Radiologists (ACR) [9,17], showed that the best cutoff value was TI-RADS grade 4 (Figure 7). The accuracy, sensitivity, specificity, PPV, NPV, and AUC of the junior radiologist, intermediate-level radiologist, and senior radiologist are shown in Table 2. Comparison of the performance of the CAD system, radiologists at all levels, and the junior radiologist combined with CAD for thyroid cancer diagnosis are shown in Table 3.

The accuracy, sensitivity, specificity, and AUC of the CAD system for the diagnosis of thyroid nodules of different sizes

The accuracy, sensitivity, specificity, and AUC of the CAD system for the diagnosis of small thyroid nodules were 85.71%, 88.68%, 83.15%, and 0.88, respectively. The accuracy, sensitivity, specificity, and AUC of the CAD system for the diagnosis of medium-sized thyroid nodules were 84.17%, 91.18%, 79.62%, and 0.92, respectively. The accuracy, sensitivity, specificity, and AUC of the CAD system for the diagnosis of large thyroid nodules were 66.96%, 50.00%, 89.79%, and 0.82, respectively (Table 4).
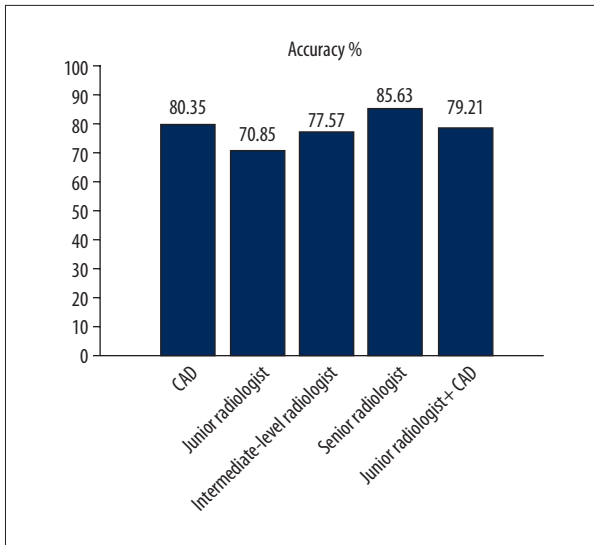
This work is licensed under Creative Common Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

e918452-6

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Figure 4. The accuracy of the computer-aided diagnosis (CAD) system. Senior, intermediate-level, and the junior radiologist combined with the CAD system for the classification of thyroid lesions as benign or malignant and the Thyroid Imaging Reporting and Data System (TI-RADS) categories. CAD – computer-aided diagnosis system; TI-RADS – Thyroid Imaging Reporting and Data System.

Figure 6. The specificity of the computer-aided diagnosis (CAD) system. Senior, intermediate-level, and the junior radiologist combined with the CAD system for the classification of thyroid lesions as benign or malignant and the Thyroid Imaging Reporting and Data System (TI-RADS) categories. CAD – computer-aided diagnosis system; TI-RADS – Thyroid Imaging Reporting and Data System.
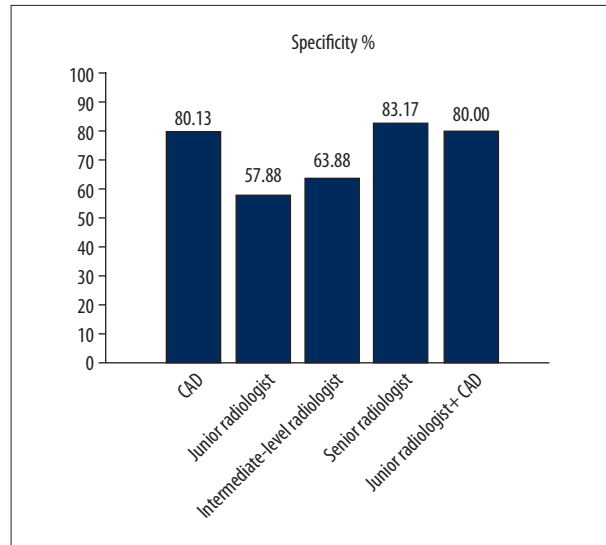

Figure 5. The sensitivity of the computer-aided diagnosis (CAD) system. Senior, intermediate-level, and the junior radiologist combined with the CAD system for the classification of thyroid lesions as benign or malignant and the Thyroid Imaging Reporting and Data System (TI-RADS) categories. CAD – computer-aided diagnosis system; TI-RADS – Thyroid Imaging Reporting and Data System.
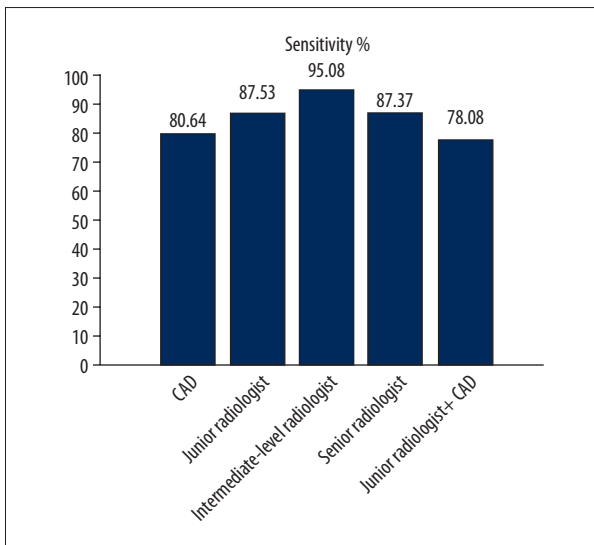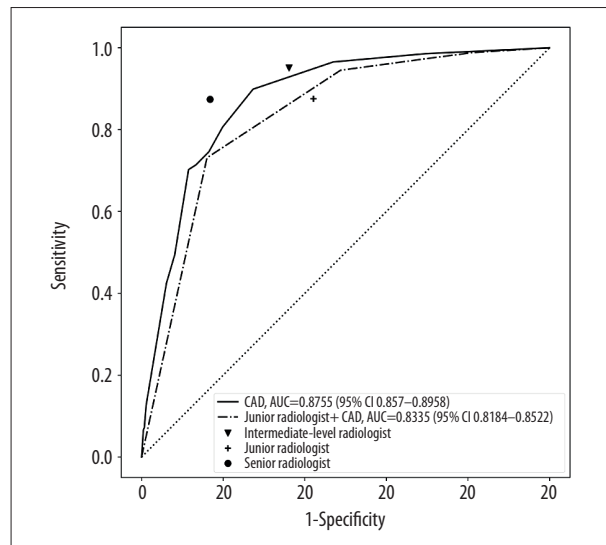

Figure 7. The receiver operating characteristic (ROC) curves for the performance of the computer-aided diagnosis (CAD) system and junior radiologist combined with the CAD system on the classification of thyroid nodules.

Table 2. The diagnosis of the thyroid nodules by the computer-aided diagnosis (CAD) system, radiologists at different levels, and the junior radiologist using the CAD system.

	Accuracy %	Sensitivity %	Specificity %	PPV %	NPV %	AUC
CAD	80.35	80.64	80.13	76.02	84.12	0.87
Junior radiologist	70.85	87.53	57.88	61.76	85.67	0.73
Intermediate-level radiologist	77.57	95.08	63.88	67.28	94.33	0.80
Senior radiologist	85.63	87.37	83.17	87.97	83.38	0.91
Junior radiologist+CAD	79.21	78.08	80.00	73.12	83.97	0.83

CAD – computer-aided diagnosis; PPV – positive predictive value; NPV – negative predictive value; AUC – area under the curve.

Table 3. Comparison of the computer-aided diagnosis (CAD) system, radiologists at different levels, and the junior radiologist using the CAD system in terms of diagnostic performance of the thyroid nodules.

	P-value		
	Accuracy	Sensitivity	Specificity
CAD *vs.* junior radiologist	<0.01	0.013	<0.01
CAD *vs.* intermediate-level radiologist	0.174	<0.01	<0.01
CAD *vs.* senior radiologist	0.005	0.008	0.274
CAD *vs.* junior radiologist+CAD	0.573	0.415	0.959
Junior radiologist+CAD *vs.* junior radiologist	<0.01	0.001	<0.01
Junior radiologist+CAD *vs.* intermediate-level radiologist	0.427	<0.01	<0.01
Junior radiologist+CAD *vs.* senior radiologist	0.001	0.001	0.25

CAD – computer-aided diagnosis.

Table 4. Diagnostic performance of the computer-aided diagnosis (CAD) system in differentiating the size of different thyroid nodules.

	Accuracy %	Sensitivity %	Specificity %	AUC
<1 cm	85.71	88.68	83.15	0.88
1–2 cm	84.17	91.18	79.62	0.92
≥2 cm	66.96	50.00	89.79	0.82

CAD – computer-aided diagnosis; AUC – area under the curve.

Table 5. Comparison of the computer-aided diagnosis (CAD) system in differentiating the size of different thyroid nodules.

	P-value		
	Accuracy	Sensitivity	Specificity
<1 cm *vs.* 1–2 cm	0.851	0.417	0.537
<1 cm *vs.* ≥2 cm	0.003	<0.01	0.057
1–2 cm *vs.* ≥2 cm	0.005	<0.01	0.006

CAD – computer-aided diagnosis.

Comparison of the diagnostic performance of the CAD system for the diagnosis of thyroid nodules of different sizes

The accuracy, sensitivity, and specificity of the CAD system for the diagnosis of small thyroid nodules diagnosis did not differ from those of medium-sized thyroid nodules (P=0.851; P=0.417; P=0.537). The accuracy and sensitivity of the CAD system for the diagnosis of large thyroid nodules were less than for small thyroid nodules (P=0.003; P<0.01). The specificity of the CAD system in the diagnosis of large thyroid nodules was not significantly different from that of small thyroid nodules (P=0.057). The accuracy and sensitivity of the CAD system for the diagnosis of large thyroid nodules were significantly less than for medium-sized thyroid nodules (P=0.005; P<0.01). The specificity of the CAD system for the diagnosis of large thyroid nodules was significantly greater than for medium-sized thyroid nodules (P=0.006) (Table 5).

Discussion

In this study, the role of the computer-aided diagnosis (CAD) system in the ultrasound diagnosis of thyroid cancer was evaluated, with emphasis on its added value for the junior radiologist. This retrospective study showed that when compared with the definitive histopathology diagnosis, the CAD system for the diagnosis of thyroid cancer had an accuracy of 80.35%, a sensitivity of 80.64%, a specificity of 80.13%, a positive predictive value (PPV) of 76.02%, a negative predictive value (NPV) of 84.12%, and an area under the receiver operating characteristic (ROC) curve (AUC) of 0.87. For the junior radiologist and the combined use of the CAD system, the accuracy, sensitivity, specificity, PPV, NPV, and AUC for thyroid cancer diagnosis were 79.21%, 78.08%, 80.00%, 73.12%, 83.97%, and 0.83 respectively.

In clinical practice, radiologists use current clinical diagnostic guidelines, including those from the American Thyroid Association (ATA) and the American College of Radiology (ACR) to identify thyroid nodules. Based on the Thyroid Imaging Reporting and Data System (TI-RADS) scores [9,17], radiologists may use the CAD system, as the imaging categories on ultrasound can be digitized and used in the CAD system learning program. Therefore, an automatic or semi-automatic classification system based on imaging features is possible and can be based on the recommended TI-RADS reporting scores and classification. In the present study, thyroid ultrasound images were generated by different types of ultrasound equipment, which helped to increase data diversity, develop training algorithms, and test the radiologist's diagnostic ability.

Several studies have reported that the diagnostic performance of CAD systems is comparable to that of experienced radiologists. The sensitivity of the CAD system for the diagnosis of thyroid cancer was previously reported to be similar to that of an experienced radiologist, but the diagnosis by the experienced radiologist had a higher specificity when compared with CAD [15,16]. However, Wang et al. [11] reported that the accuracy, sensitivity, PPV, and NPV of an artificial intelligence (AI) diagnosis system for malignant thyroid nodules diagnosis were 90.31%, 90.5%, 95.22%, and 80.99%, respectively, and the performance of the AI diagnosis system was not significantly different from that of an experienced radiologist. In this previous study, the experienced radiologist had a lower diagnostic specificity when compared with the AI system (77.98% vs. 89.91%) [11].

Although the findings from the present study demonstrated the potential usefulness of the CAD system, this was a small and preliminary study, and these findings should not be generalized without further validation. Also, the results were acquired only through comparison with experienced radiologists. The usefulness of the CAD system for diagnostic thyroid ultrasound may be different depending on the range of different experiences diagnostic thyroid ultrasound in clinical practice. This study showed that when compared with the junior radiologist, the CAD system resulted in increased accuracy and specificity and lower sensitivity in the classification of thyroid cancer. Also, the CAD system showed increased accuracy and specificity and lower sensitivity than the intermediate-level radiologist, and lower accuracy, sensitivity, and specificity than the senior radiologist in the classification of thyroid cancer.

Several previously reported studies have demonstrated the application of different types of CAD systems for thyroid ultrasound [11,18,19]. These studies have reported that the CAD systems improved the diagnostic performance of thyroid ultrasound [11,18,19]. The findings from the present study concluded that the CAD system was helpful to the junior radiologist, and resulted in significant improvement in the diagnostic performance. With the help of the CAD system, the diagnostic performance of the junior radiologist was similar to that of the intermediate-level radiologist. Also, the use of the CAD system improved the diagnosis of thyroid cancer by the junior radiologist, as the accuracy increased from 70.85% to 79.21%, the sensitivity decreased from 87.53% to 78.08%, and the specificity increased from 57.88% to 80.00%. Also, the AUC of 0.78 increased to 0.83, which exceeded the intermediate-level radiologist's AUC of 0.80. The accuracy, specificity, PPV, and AUC of the junior radiologist were improved, which supported the view that the CAD system may have a potential role in providing a supplementary, confirmatory, or a second opinion that can be used to plan the next stage in the management of the patient with a thyroid nodule.

However, although the use of the CAD system may enable improvement of the accuracy, specificity, PPV, and AUC of the diagnostic ability of the junior radiologist, as shown in this study, its use may be accompanied by reduced sensitivity and NPV, which must be considered when using the CAD system in clinical practice. Therefore, from the findings of this study, the CAD system might only be recommended as an adjunctive diagnostic method for radiologists rather than a single or definitive diagnostic method in patients with a thyroid nodule, even for junior radiologists.

Subgroup stratification analysis was performed to investigate whether different sizes of thyroid nodules affected the performances of the CAD system. Subgroup analysis showed that the diagnostic performance of the thyroid ultrasound CAD system in distinguishing between the sizes of the different thyroid nodules was not consistent in each group. The diagnostic performance of the CAD system in identifying large thyroid nodules was less than that of small thyroid nodules and medium-sized thyroid nodules. The reason for this finding might be that large thyroid nodules have a boundary that is beyond the thyroid gland, which makes it impossible for the CAD system to distinguish between the nodules and the normal thyroid gland, resulting in poor performance in the diagnosis of the large thyroid nodule by the CAD system.

In this study, a total of 29.7% (234/789) thyroid nodule pathological results were confirmed by core needle biopsy (CNB). Before CNB or fine-needle aspiration cytology (FNAC), different biopsy methods were chosen based on the ultrasound characteristics of the thyroid nodules. For example, nodules with calcification, diffuse microcalcification in the thyroid gland, nodules with a maximum diameter of more than 2 cm, and when the cytology results of the thyroid nodules remained unclear after FNAC. In these cases, CNB was performed for most patients.

This study had several strengths that contribute to improving clinical practice. First, the CAD system in this study can automatically analyze the ultrasound features of thyroid nodules and then give their TI-RADS grades. This property of the CAD system enables radiologists to know how the system identifies thyroid cancer. Artificial intelligence (AI) is often regarded as impractical and unreliable but, as this study has shown, the CAD system can automatically provide TI-RADS grades of thyroid nodules, which demonstrates an opportunity for the combination between clinician and machine in clinical practice. Also, in the present study, the CAD system compared the performance of radiologists with different diagnostic experience, which could be tested further in clinical practice. This study assessed the added diagnostic value that the CAD system provided to the junior radiologist. Lastly, the diagnostic performance of the CAD system for thyroid nodules of different sizes was evaluated, which was able to reflect the diagnostic performance of the CAD system further.

This study also had several limitations. This was a retrospective study at a single center in which the study was conducted, and the results were analyzed by the authors, all of which could have introduced study bias. All the data analyzed were derived from static images, which may have resulted in misinterpretation. Patients with symptoms of thyroiditis and suspicious thyroid nodules with inflammation were not included in the study. The use of FNAC of the thyroid has a false-positive and false-negative rate of up to 5%, which may have affected the results. Also, the study only assessed the adjunctive role of the CAD system for the junior radiologist but did not assess the role of the CAD system for intermediate-level radiologists and senior radiologists. This study included a single radiologist in each group, and future studies should include a larger study group. Finally, the CAD system requires more training, development, and practical evaluation to improve its diagnostic performance.

Conclusions

This study aimed to develop a computer-aided diagnosis (CAD) system based on the Thyroid Imaging Reporting and Data System (TI-RADS) to distinguish benign from malignant thyroid nodules by analyzing ultrasound images. The proposed CAD system, based on the TI-RADS, has the potential to improve the diagnostic performance of junior radiologists in examining thyroid nodules using ultrasound.

References:

1. Haugen BR, Alexander EK, Bible KC et al: 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association Guidelines Task Force on thyroid nodules and differentiated thyroid cancer. Thyroid, 2016; 26: 1–133

2. Brito JP, Morris JC, Montori VM: Thyroid cancer: Zealous imaging has increased detection and treatment of low risk tumours. BMJ, 2013; 347: f4706

3. Miller KD, Siegel RL, Lin CC et al: Cancer treatment and survivorship statistics, 2016. Cancer J Clin, 2016; 66(4): 271–89

4. Grani G, Lamartina L, Cantisani V et al: Interobserver agreement of various thyroid imaging reporting and data systems. Endocr Connect, 2018; 7: 1–7

5. Grant EG, Tessler FN, Hoang JK et al: Thyroid ultrasound reporting lexicon: White paper of the ACR Thyroid Imaging, Reporting and Data System (TI-RADS) Committee. J Am Coll Radiol, 2015; 12: 1272–79

6. Kwak JY, Han KH, Yoon JH et al: Thyroid imaging reporting and data system for US features of nodules: A step in establishing better stratification of cancer risk. Radiology, 2011; 260: 892–99

7. Cappelli C, Castellano M, Pirola I et al: The predictive value of ultrasound findings in the management of thyroid nodules. QJM, 2007; 100(1): 29–35

8. Horvath E, Silva CF, Majlis S et al: Prospective validation of the ultrasound based TI-RADS (Thyroid Imaging Reporting And Data System) classification: Results in surgically resected thyroid nodules. Eur Radiol, 2017; 27: 2619–28

9. Tessler FN, Middleton WD, Grant EG et al: ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White paper of the ACR TI-RADS Committee. J Am Coll Radiol, 2017; 14: 587–95

10. Gitto S, Grassi G, De Angelis C et al: A computer-aided diagnosis system for the assessment and characterization of low-to-high suspicion thyroid nodules on ultrasound. Radiol Med, 2019; 124(2): 118–25

11. Wang L, Yang S, Yang S et al: Automatic thyroid nodule recognition and diagnosis in ultrasound imaging with the YOLOv2 neural network. World J Surg Oncol, 2019; 17: 12

12. Acharya UR, Faust O, Sree SV et al: ThyroScreen system: High resolution ultrasound thyroid image characterization into benign and malignant classes using novel combination of texture and discrete wavelet transform. Comput Methods Programs Biomed, 2012; 107: 233–41

13. Acharya UR, Vinitha Sree S, Krishnan MM et al: Non-invasive automated 3D thyroid lesion classification in ultrasound: A class of ThyroScan systems. Ultrasonics, 2012; 52: 508–20

14. Chang Y, Paul AK, Kim N et al: Computer-aided diagnosis for classifying benign versus malignant thyroid nodules based on ultrasound images: A comparison with radiologist-based assessments. Med Phys, 2016; 43: 554

15. Choi Y J, Baek JH, Park HS et al: A computer-aided diagnosis system using artificial intelligence for the diagnosis and characterization of thyroid nodules on ultrasound: Initial clinical assessment. Thyroid, 2017; 27: 546–52

16. Gao L, Liu R, Jiang Y et al: Computer-aided system for diagnosing thyroid nodules on ultrasound: A comparison with radiologist-based clinical assessments. Head Neck, 2018; 40(4): 778–83

17. Tessler FN, Middleton WD, Grant EG, Hoang JK: Re: ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White paper of the ACR TI-RADS Committee. J Am Coll Radiol, 2018; 15: 381–82

18. Jeong EY, Kim HL, Ha EJ et al: Computer-aided diagnosis system for thyroid nodules on ultrasonography: Diagnostic performance and reproducibility based on the experience level of operators. Eur Radiol, 2019; 29: 1978–85

19. Yoo YJ, Ha EJ, Cho YJ et al: Computer-aided diagnosis of thyroid nodules via ultrasonography: Initial clinical experience. Korean J Radiol, 2018; 19: 665–72