(Austin Publishing Group

## **Research Article**

# TIRADS-Based Artificial Intelligence Systems for Ultrasound Imaging of Thyroid Nodules: A Systematic Review

## Sharifi Y<sup>1\*</sup>, Danay Ashgzari M<sup>2</sup>, Naseri Z<sup>3</sup> and Amiri Tehranizadeh A<sup>3</sup>

<sup>1</sup>Department of Medical Informatics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran <sup>2</sup>Department of Computer, Faculty of Engineering, Islamic Azad University of Mashhad, Mashhad, Iran <sup>3</sup>Department of Medical Informatics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

#### \*Corresponding author: Sharifi Y,

Department of Medical Informatics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran **Tel:** 00989153523360; **Email:** sharifiy@mums.ac.ir

**Received:** March 15, 2025 **Accepted:** April 02, 2025 **Published:** April 04, 2025

#### Abstract

**Objective**: The Thyroid Imaging Reporting and Data Systems (TI-RADS) is a standard terminology that classifies thyroid nodules according to their potential risk of cancer to reduce unnecessary biopsies, minimize variations in interpreting thyroid nodule images, and improve diagnostic accuracy. This study aims to comprehensively review articles that utilize AI techniques to develop decision support systems for analyzing ultrasound images of thyroid nodules, following different TIRADS guidelines.

**Materials and Methods**: We followed a five-step process, this included identifying the key research questions, outlining the literature search strategies, establishing criteria for including and excluding studies, assessing the quality of the studies, and extracting the relevant data. We created a comprehensive search string to gather all relevant English-language studies up to January 2024 from the PubMed, Scopus, and Web of Science databases, and we also followed the PRISMA diagram.

**Results:** In this review, forty-four papers were included, and the most important properties of these papers, including dataset characteristics, AI technical specifications, results and outcome metrics, metrics, limitations, and contributions, were extracted.

**Conclusion:** We evaluated the technical characteristics and various aspects used in the development of artificial intelligence CAD systems based on various TI-RADS. This review demonstrates that AI advancements, especially deep learning methods, have significantly enhanced CAD systems for evaluating thyroid nodules. However, comprehensive datasets, multimodal images, and standard evaluation metrics are needed to further enhance machine learning models. Our study aims to provide researchers and physicians with a summary of the current advancements in this field to guide future investigations.

**Keyword:** Thyroid nodules; Artificial Intelligence; TIRADS; Ultrasonography; Computer-assisted diagnosis

## Introduction

The thyroid gland, a small yet crucial endocrine organ situated in the anterior aspect of the neck, plays a significant role in the regulation of metabolism and various bodily functions [1]. Thyroid nodules are frequently encountered in clinical practice, with the majority of cases being benign. However, accurately differentiating between benign and malignant nodules to guide appropriate management strategies is paramount. The evaluation of thyroid nodules often involves a combination of clinical assessment, imaging studies such as ultrasound, and fine needle aspiration biopsy for cytological examination [2].

Fine needle aspiration (FNA) is an invasive procedure used to evaluate thyroid nodules for the presence of cancerous cells. However, it is common practice for many nodules to undergo a biopsy to identify a small percentage of cases that may be malignant. It is important to consider the potential burden that FNA procedures can place on healthcare systems, as they can result in significant costs and create stress and anxiety for patients. Therefore, it is crucial for healthcare providers to carefully evaluate the necessity of such procedures and consider alternative approaches when possible [3,4]. Thyroid ultrasound imaging plays a crucial role in the identification of thyroid nodules because of its accessibility, noninvasive nature, and cost effectiveness. This procedure allows clinicians to visualize the thyroid gland and any abnormalities present within it [5]. Furthermore, it is a safe and convenient diagnostic tool that can be easily performed in outpatient settings, making it a valuable resource for monitoring thyroid health and guiding treatment decisions [5].

The Thyroid Imaging, Reporting, and Data System (TI-RADS) was established to provide a standardized framework for categorizing thyroid nodules according to their specific characteristics associated with risk. This system aims to mitigate issues surrounding the

variability and low reproducibility that often arise in the detection and interpretation of nodule features among different physicians [6]. By implementing TI-RADS, healthcare providers can ensure a more consistent and reliable approach to evaluating thyroid nodules, ultimately leading to more accurate diagnoses and treatment decisions for patients [7]. There are several variations of TIRADS, each with its own specific criteria and scoring system. These variations, such as the American College of Radiology (ACR) TIRADS [8], the Korean Society of Thyroid Radiology (KTIRADS) [9], ACE [10], ATA [11], Kwak-TIRADS [12], and the European Thyroid Imaging and Reporting System (EU-TIRADS) [13], aim to standardize the interpretation and management of thyroid nodules. AI-based approaches, such as machine learning and deep learning algorithms, have demonstrated significant potential in enhancing the accuracy and efficacy of thyroid nodule evaluation. These advancements not only help reduce variability among observers but also contribute to improving diagnostic outcomes by identifying patterns and trends that may not be easily identifiable by clinicians alone [14-18].

With advancements in medical technology, computer-aided detection (CAD) systems have been developed to assist radiologists in analyzing ultrasound images of thyroid nodules. These CAD systems can help in the early detection of suspicious nodules, leading to timely intervention and improved patient outcomes. By combining the expertise of radiologists with the efficiency and accuracy of CAD systems, healthcare professionals, by minimizing the subjective nature of traditional diagnostic methods, can provide more precise and reliable diagnoses and treatment plans for patients with thyroid nodules [18-20].

The development of AI-driven TIRADS models, which combine computerized analysis of ultrasound images with established risk stratification systems, represents a progressive step in the field of thyroid imaging [14,21,22].

The classification of thyroid nodules via various TIRADS systems has been the subject of several studies, highlighting the importance of evaluating these systems in depth. The primary objective of this study is to explore the utilization of artificial intelligence CAD systems in the ultrasound image classification of thyroid nodules via various TIRADS systems. It is crucial to consider factors such as dataset characteristics, technical specifications of the network, evaluation metrics, results, advantages, obstacles, and limitations.

By analyzing the literature, this research aims to offer a comprehensive understanding of the role of AI techniques in the development of TIRADS-based decision support systems for this purpose to highlight the challenges and prospects that lie ahead in the integration of these groundbreaking technologies into clinical settings. However, to the best of our knowledge, no systematic review has explicitly focused on this field. The insights gained from this study could serve as a valuable resource for researchers and developers looking to create more effective systems with improved efficiency. Ultimately, the implementation of these systems could help reduce unnecessary thyroid nodule biopsies, address issues of over care, enhance the reproducibility and reliability of ultrasound diagnostics, and provide educational support for less experienced physicians.

To carry out these tasks, the following research questions are

proposed to direct this systematic literature review:

- What is the best artificial intelligence technique for implementing a thyroid nodule classification system based on TI-RADS?

• What is the size of the appropriate dataset for the successful implementation of these systems?

• What are the most common neural network architectures used in these systems?

• What is the most common TI-RADS used in these systems?

• What are the most common evaluation metrics in these systems?

• What are the limitations and future directions in this field?

The remainder of this paper is structured as follows: Section 2 outlines the methodology of the systematic review. Section 3 details the findings of various uses of AI systems based on TIRADS on ultrasound images of thyroid nodules. Finally, a discussion will be presented, and conclusions and future works will be drawn.

## **Materials and Methods**

This systematic review involves five main steps: literature search, study selection, study quality assessment, data extraction, and analysis. Further details of each step are presented in the subsequent subsections. Importantly, the protocol for this systematic review was registered in the PROSPERO database in August 2024 [23]. (Registration number: CRD42024551311).

## Literature Search

This study conducted a systematic review to retrieve all relevant English language articles up to January 2024 via the PubMed, Scopus, and Web of Sciences databases. The search terms included "Thyroid Imaging Reporting and Data System", "Artificial Intelligence", "ultrasonography" and their related terms (Table 1). In addition, the Medical Subject Headings (MeSH) vocabulary and synonym keywords were utilized.

## **Study Selection**

This study adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines [24]. To select relevant articles, we defined the inclusion and exclusion criteria.

The inclusion criteria were as follows:

1- Articles that have implemented an AI system based on TIRADS.

2- Articles on ultrasound imaging of the thyroid.

The exclusion criteria were as follows:

1. Nonoriginal articles such as review articles, comments, and editorials.

2. Conference abstracts and unpublished articles.

3. Articles that do not use TIRADS.

4. Articles that do not include ultrasound images.

5. Articles that evaluate existing AI systems on the basis of TIRADS.

## **Study Quality Assessment**

The included studies were subjected to a quality assessment process to evaluate the credibility and strength of the articles. We used a modified quality assessment with 13 questions and three options ("Yes"= 1, "Partly"= 0.5, and "No"= 0), as suggested by Sharifi et al. [18] (Table 1S).

Two independent researchers with backgrounds in systematic review, machine learning, deep learning, and medical informatics evaluated the quality of the included studies and resolved any discrepancies in their findings by consulting a third researcher to reach a unanimous conclusion.

## **Data Extraction**

Two reviewers independently evaluated and extracted data from the included articles, using a predesigned table in Microsoft Excel to ensure accuracy. A pilot test was subsequently conducted on twenty random studies to confirm the reliability of their data extraction. The calculated kappa statistic [25] indicated strong agreement in data interpretation (kappa statistic = 0.85). The following major aspects of the included studies were extracted: paper information, patient information, dataset characteristics, technical specifications, results, outcome metrics, limitations, and contributions.

## **Data Analysis**

In this section, the major aspects of the articles that implemented a TIRADS-based AI system are analyzed.

## **Results**

#### Literature Search and Study Selection

The identification of potentially related articles to TIRADS-based artificial intelligence systems on US images of thyroid nodules in this systematic review adheres to the PRISMA flow diagram and guidelines [24]. Figure 1 displays the PRISMA diagram for this study, which comprises four primary phases. The initial phase involved identifying relevant English language articles via the PubMed, Scopus, and Web of Sciences databases until January 2024, on the basis of the search strategy outlined in Section 2-1. Initially, 618 papers were found, and after removing duplicates, 521 papers remained. In the next stage, after screening the titles and abstracts, 443 unsuitable articles were removed, leaving 88 articles for further consideration. In the third phase, we evaluated the suitability of the articles by reading the full texts and applying the inclusion and exclusion criteria outlined in Section 2-2. As a result, 44 articles were eliminated from the study. In the fourth phase, 44 articles were chosen for additional qualitative analysis.

## **Literature Sources**

The analysis involved reviewing 44 selected articles to investigate TIRADS-based artificial intelligence research on ultrasound images of thyroid nodules.

These articles were published from 2017--2024. They were categorized as follows: Q1 (66%), Q2 (27%), and Q3 (7%). Summary details regarding these articles can be found in **Table 2S** of the Supplementary data.

#### **Austin Publishing Group**



## Study Quality Assessment

To evaluate the quality of the selected articles, two independent researchers responded to 13 quality answers [18] for articles that implemented a TIRADS-based AI system as previously stated. If there were any discrepancies in their evaluations, they sought advice from a third researcher. The final scores were subsequently calculated by adding the scores of these answers for each article that could receive a score ranging from zero to 13.

Furthermore, the articles are divided into three groups on the basis of their scores, namely, "low-score," "mid-score," and "high-score", by splitting the score range into three equally sized intervals: [0 - 4.33), [4.33 - 8.66), and [8.66 - 13], respectively.

The details and results of the quality questionnaire are shown in Table 3S in the supplementary data. According to the computed scores, the articles are distributed as follows: 2% low-score, 17% midscore, and 81% high-score categories.





#### **Data Analysis**

In this review, forty-four papers used TIRADS in the implementation of an AI system for analyzing ultrasound images of thyroid nodules.

Figure 1S in the supplementary data illustrates the number of articles by year until January 2024, and the most important properties of these papers, including dataset characteristics, AI technical specifications, results and outcome metrics, limitations, and contributions, are presented in Table 2, Table 3, and Table 4.

## **Discussion**

The primary objective of this study was to conduct a systematic review of articles related to TIRADS-based CAD systems for analyzing thyroid ultrasound images.

Among the initial 618 publications, 44 articles published up to January 2024 were selected for this study. As depicted in Figure 1S, articles in this field have been published from 2017 to 2024, with the highest number of publications in 2023 (n=14, 30.4%). All the articles that have utilized TIRADS guidelines to implement AI systems for analyzing thyroid ultrasound images have focused on classification tasks. The expected growth in this field is due to CAD systems being designed to detect suspicious nodules and differentiate between benign and malignant nodules in thyroid ultrasound images.

## **Comparison of Outcome Metrics**

The use of various evaluation metrics in these studies makes it difficult to assess and compare the performance of the CAD systems being presented. In these studies, as depicted in Figure 2S in supplementary data, the most popular metrics are accuracy (n=35, 21%), sensitivity (n=34, 20%), specificity (n=34, 20%), area under the curve (AUC) (n=23, 13%), PPV (n=18, 10%), NPV (n=18, 4011%), and F score (n=8, 5%).

## **Dataset Comparison**

The performance of research articles has been validated via various datasets of different sizes and types, including local and public datasets.

Table 4S and Figure 2 present statistical information about the size and type of datasets included in the studies. The used dataset

consisted of a minimum of 134 images from a local source and a maximum of 31888 images, which included both a local dataset and a public dataset. In Figure 2, it is clear that only a small fraction (n:3, 4.4%) of the studies use public datasets, making it difficult to compare their methods.

The public ultrasound thyroid datasets used in these papers are the Thyroid Digital Image Database (TDID), provided by Pedraza et al. [74], and an open-source dataset from the scientific community [75]. Among the studies that used local datasets, (n=31, 68.9%) utilized one dataset (one center), and (n=9, 20%) utilized two to four datasets (multiple centers).

## Image Preprocessing and Augmentation

Image preprocessing is a crucial step in medical image analysis. It sets the foundation for accurate image interpretation and insight extraction. This phase often includes detailed operations such as cropping and resizing the region of interest (ROI), which are essential for focusing the analysis on the most relevant aspects of the image [29,32,36-38,44].

Fundamental to image preprocessing are processes such as binarization, which effectively separates objects from their backgrounds, and normalization, which is crucial for ensuring that intensity values remain consistent across a dataset, facilitating more reliable comparisons and evaluations [30,31,53].

In addition to these core techniques, various image filtering methodologies, such as median filters [34] and bilateral filtering [65], are used to reduce noise and enhance important features in images, thus improving the clarity and usefulness of visual data. Specialized preprocessing techniques, such as removing patient identification details and any misleading markers (artifacts) from nodules [45,64,66,68], improve image clarity, enabling more accurate analysis and diagnosis. This is often complemented by additional image enhancement techniques and advanced denoising strategies. All of these techniques aim to improve the overall quality of the images. Such comprehensive preprocessing efforts are critical, as they significantly increase the reliability and accuracy of image-based assessments across a multitude of applications. This informs decision-making processes and enhances the efficacy of subsequent analyses [68].

In addition to the initial processing steps, detection of the region of interest (ROI) is conducted to identify and isolate the relevant areas within the images. The images are subsequently resized to standardized dimensions to ensure uniformity. A manual cropping process is employed to format the images into a square shape, facilitating consistency across the dataset and enhancing the effectiveness of subsequent analysis [44,45]. The detection of regions of interest (ROIs) enhances the accuracy of diagnostic assessments by prioritizing specific areas within an image that warrant further examination. Many studies have employed manual methodologies [27,30,33,36,37,44,45,52-54,58,59,61]. These manual techniques, while traditional, often require considerable time and are subject to human error. This has prompted a shift toward more automated approaches. In contrast, a few research endeavors have embraced automated methods for detecting regions of interest (ROIs) [31,48,66]. With the continuous advancement of technology, there is a growing

opportunity to improve manual techniques by integrating automated solutions. This approach has the potential to increase the efficiency and consistency of nodule detection.

Among modern techniques, deep learning models, such as RetinaNet [32] and Faster R-CNN [42], which utilize cutting-edge frameworks, have gained significant attention because of their ability to enhance detection capabilities. In terms of segmentation, various methodologies have emerged, including StableSeg [28], U-Net++ [47], U-Net [49], and deep learning-based segmentation approaches such as SkaNet [67] and EfficientNet B6 [68].

Moreover, various tools and applications for detecting regions of interest (ROIs) and facilitating their extraction and analysis are discussed in the literature. One notable tool is ePADlite, which is a semiautomated segmentation tool integrated within the Electronic Physician Annotation Device [35]. Manual ROI tools such as ITK-SNAP [38,50], MATLAB [51], and ImageJ software (version 1.48, National Institutes of Health, USA) serve as alternative options for researchers. Additionally, LabelMe software [65] is recognized for its ability to facilitate precise annotation and segmentation tasks within this dynamic field of study.

Medical image augmentation plays a vital role in overcoming the challenges associated with limited medical image datasets. Artificially expanding the volume and diversity of training data through augmentation techniques such as rotation, flipping, zooming, mirroring, shifting, scaling and adjusting brightness or contrast [27,30,32,33,54,60] or adding Gaussian noise [30,33,45] can significantly enhance the performance and robustness of machine learning models.

Transfer learning has become a highly effective strategy for addressing the challenges of insufficient medical imaging data and improving generalizability across various applications. Several studies have shown that using pretrained models significantly enhances the performance of machine learning frameworks, providing a strong starting point for tasks where data scarcity is a concern. Many researchers in deep learning have specifically utilized various architectures pretrained on ImageNet datasets, demonstrating the adaptability of these models to medical imaging tasks [27-33,35-37,44,54,55,60,65,66,68]. This not only increases the prediction accuracy but also speeds up the training process, ultimately leading to better clinical outcomes.

Compared with traditional machine learning methods, contemporary deep learning methods tend to use image augmentation, nodule detection, and segmentation techniques. The ratio of these approaches used is illustrated in Figure 3.

## **Detailed Technical Comparison**

Among the reviewed articles, two distinct approaches were identified: traditional machine learning and deep learning methodologies. Recent studies have shown that a variety of traditional machine learning techniques have been used to analyze medical images. Several articles have used XGBoost to improve the accuracy of predicting and classifying ACR-TIRADS features. XGBoost is an advanced machine learning algorithm that uses a decision treebased framework to efficiently process and analyze complex datasets, thereby significantly improving diagnostic performance in thyroid imaging assessments [39,50,52,62,75]. In several articles, researchers have explored and compared multiple machine learning approaches to identify the most effective method for specific applications. Some frequently used techniques include support vector machines (SVMs), artificial neural networks (ANNs), logistic regression (LR), K-nearest neighbors (KNNs), and random forests (RFs) [30,34,39,45,46,52,5 9,61,62,69]. Through systematic evaluation, the goal is to select the optimal algorithm that provides the best performance on the basis of the data characteristics and desired outcomes.

Deep learning methods stand out from traditional machine learning approaches, primarily because they rely on transfer learning, especially in the medical imaging field, where labeled data are often limited. As shown in the previous section, out of 19 studies that used transfer learning, the majority (17 out of 19) employed architectures that were pretrained on ImageNet.

This trend demonstrates the effectiveness of ImageNet as a fundamental dataset, providing strong feature extraction capabilities that improve performance across various tasks.

The detection and segmentation of nodules have become important parts of deep learning frameworks. In their work, researchers used StableSeg [27] because of its strong performance in segmentation tasks, whereas Unet [49] and Unet++ [47] were chosen for their excellent capabilities in segmenting biomedical images. With respect to architecture, researchers have relied mainly on previously developed methods to improve their models, including ResNet50 [28,29,37,44,47,55,66], DenseNet121 [29,30,32,44,56,68], InceptionV3 [30], MobileNetV2 [33,35], VGG16 [36], EfficientNetB7 [45], EfficientNetB3 [65], InceptionResNetV2 [60], and GoogleNet [44].

In the machine learning literature, significant results were reported by Vadhiraj et al. in [34]. They used a median filter and image binarization techniques for effective image preprocessing and segmentation. Additionally, they utilized the gray-level co-occurrence matrix (GLCM) to extract seven relevant features from ultrasound images.

The support vector machine (SVM) achieved impressive 96% accuracy in classification, as reported by Gomes et al. [58]. In their study, 27 morphological and geometric features were meticulously extracted from images obtained from a publicly available U.S. thyroid nodule image database. These features were then analyzed via an RF classifier (RFC), which achieved a remarkable classification accuracy of 99.33%.

Yu et al. proposed an innovative fused deep learning model for analyzing thyroid nodules. The model begins by extracting 33 clinically significant statistical features. After conducting principal component analysis (PCA) for dimensionality reduction, the top four features are integrated with the feature map generated by EfficientNet B3 [65]. This approach achieved an impressive accuracy rate of 96.4%, demonstrating its effectiveness in the given application.

Among the articles included, Figure 3S in supplementary data, shows that the most commonly implemented TI-RADS guideline as an artificial intelligence system is the ACR-TIRADS (n=29, 61.7%).

This finding highlights the high effectiveness of this ultrasound risk stratification system.

## **Challenges and Future Perspectives**

In machine learning (ML) and deep learning (DL), many studies face significant challenges due to inadequate and unbalanced datasets. These issues are largely attributed to local data scarcity and improper labeling of features by domain experts.

This issue is particularly pronounced in thyroid research, where the availability of comprehensive datasets accompanied by reliable ground truths and annotations is essential for enhancing model accuracy and robustness. A notable gap exists in the representation of benign and malignant cases across all risk levels of the TIRADS classification [76], which is crucial for improving the generalizability of AI models.

Additionally, collecting and accurately labeling thyroid nodules is a time-consuming and costly process. It requires extensive resources and specialized expertise to ensure the reliability and representativeness of the training data. Poorly labeled or insufficient data can lead to suboptimal AI outcomes, undermining the potential advantages of using AI in the diagnosis and management of thyroid conditions.

Furthermore, variability in image acquisition techniques such as multimodal images can significantly influence training results, whereas clinical decision-making often necessitates a comprehensive evaluation, including video assessments, rather than relying solely on isolated images.

Finally, we highlight important areas of future work for this line of research as follows. As medical imaging advances, it is crucial to create a comprehensive dataset from multiple centers. This dataset should contain a diverse and balanced distribution of images representing various levels of malignancy risk, along with an adequate representation of benign samples. This is essential for the progress of thyroid nodule classification. This dataset not only improves classification accuracy but also includes scoring systems based on international guidelines, such as the American College of Radiology Thyroid Imaging, Reporting and Data System (ACR-TIRADS) and its associated variants. This structured approach will be crucial in developing advanced artificial intelligence (AI) models for medical applications.

Cutting-edge image generation techniques such as generative adversarial networks (GANs) can improve the quality of images used to train AI systems in the medical imaging field. investigating the effects of different techniques to address the performance challenges associated with small datasets is necessary.

Moreover, the use of pretrained deep learning architectures designed for medical image processing is essential for improving the ability of AI systems to work with different types of medical images.

This initiative will focus not only on static images but also on evaluations of both images and videos, which will greatly enhance the decision-making capabilities of AI applications for analyzing thyroid nodules. This comprehensive approach has the potential to revolutionize the diagnostic process, providing more accurate and dependable tools for clinicians and ultimately leading to better patient outcomes.

Furthermore, the absence of uniform metrics for evaluating the effectiveness of the suggested networks complicates result comparison. Thus, it is essential to utilize standardized metrics for measuring and contrasting the performance of various methods.

## Conclusion

This article provides a comprehensive review of AI systems that utilize the Thyroid Imaging Reporting and Data System (TIRADS) for analyzing ultrasound images of thyroid nodules. This systematic review is valuable for further research and encompasses articles from reputable academic journals. The focus is on the design and development of TIRADS guidelines via machine learning (ML) methodologies, with the quality of the papers assessed on the basis of specific criteria. Notably, deep learning (DL) techniques, including transfer learning, have been widely adopted, using various segmentation and classification methods at different levels of complexity.

This review revealed that AI advancements have greatly improved computer-aided diagnosis (CAD) systems for evaluating thyroid nodules. However, the lack of comprehensive datasets, multimodal images of thyroid nodules, and standard metrics for comparison continues to hinder the progress of ML models. The results of our study are expected to aid researchers and physicians who are keen on AI-based CAD tools. This will be achieved by providing a summary of the evidence to determine the current level of advancement and guiding researchers in selecting appropriate methods for their future investigations to improve their reliability and accuracy in diagnosing thyroid-related conditions.

## Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

## References

- 1. Nilsson M, Fagman H. Development of the thyroid gland. Development. 2017; 144: 2123-2140.
- Li M, Dal Maso L, Vaccarella S. Global trends in thyroid cancer incidence and the impact of overdiagnosis. The lancet Diabetes & endocrinology. 2020; 8: 468-470.
- Acharya UR, Swapna G, Sree SV, Molinari F, Gupta S, Bardales RH, et al. A review on ultrasound-based thyroid cancer tissue characterization and automated classification. Technology in cancer research & treatment. 2014; 13: 289-301.
- Lamartina L, Deandreis D, Durante C, Filetti S. Endocrine tumours: imaging in the follow-up of differentiated thyroid cancer: current evidence and future perspectives for a risk-adapted approach. European Journal of Endocrinology. 2016; 175: R185-R202.
- Nayak R, Nawar N, Webb J, Fatemi M, Alizad A. Impact of imaging crosssection on visualization of thyroid microvessels using ultrasound: Pilot study. Scientific reports. 2020; 10: 415.
- Sharifi Y, Shafiei S, Tabesh H, Aminzadeh B, Layegh P, Mahmoodzadeh A, et al. Observation variation in ultrasonography assessment of thyroid nodules. Asia Oceania Journal of Nuclear Medicine and Biology. 2022; 10: 28.
- Yang R, Zou X, Zeng H, Zhao Y, Ma X. Comparison of diagnostic performance of five different ultrasound TI-RADS classification guidelines for thyroid nodules. Frontiers in Oncology. 2020; 10: 598225.

- Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee. Journal of the American college of radiology. 2017; 14: 587-595.
- Na DG, Baek JH, Sung JY, Kim J-H, Kim JK, Choi YJ, et al. Thyroid imaging reporting and data system risk stratification of thyroid nodules: categorization based on solidity and echogenicity. Thyroid. 2016; 26: 562-572.
- Gharib H, Papini E, Garber JR, Duick DS, Harrell RM, Hegedus L, et al. American association of clinical endocrinologists, American college of endocrinology, and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules-2016 update appendix. Endocrine practice. 2016; 22: 1-60.
- 11. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, 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.
- Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, 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: 892-899.
- Russ G, Bonnema SJ, Erdogan MF, Durante C, Ngu R, Leenhardt L. European Thyroid Association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: the EU-TIRADS. European thyroid journal. 2017; 6: 225-237.
- Chen Y, Gao Z, He Y, Mai W, Li J, Zhou M, et al. An artificial intelligence model based on ACR TI-RADS characteristics for US diagnosis of thyroid nodules. Radiology. 2022; 303: 613-619.
- Deng P, Han X, Wei X, Chang L. Automatic classification of thyroid nodules in ultrasound images using a multi-task attention network guided by clinical knowledge. Computers in Biology and Medicine. 2022; 150: 106172.
- Gu J, Xie R, Zhao Y, Zhao Z, Xu D, Ding M, et al. A machine learning-based approach to predicting the malignant and metastasis of thyroid cancer. Frontiers in Oncology. 2022; 12: 938292.
- Liu W, Lin C, Chen D, Niu L, Zhang R, Pi Z. Shape-margin knowledge augmented network for thyroid nodule segmentation and diagnosis. Computer Methods and Programs in Biomedicine. 2024; 244: 107999.
- Sharifi Y, Bakhshali MA, Dehghani T, DanaiAshgzari M, Sargolzaei M, Eslami S. Deep learning on ultrasound images of thyroid nodules. Biocybernetics and Biomedical Engineering. 2021; 41: 636-655.
- Hesson N. The Role of Computer-Aided Diagnosis (CAD) in the Diagnosis and Characterization of Thyroid Nodules when Combined with Conventional Ultrasound: A Literature Review. Journal of Medical Imaging and Radiation Sciences. 2019; 50: S15.
- Khachnaoui H, Guetari R, Khlifa N, editors. A review on deep learning in thyroid ultrasound computer-assisted diagnosis systems. 2018 IEEE International Conference on Image Processing, Applications and Systems (IPAS); 2018: IEEE.
- Han X, Chang L, Song K, Cheng L, Li M, Wei X. Multitask network for thyroid nodule diagnosis based on TI 

  RADS. Medical Physics. 2022; 49: 5064-5080.
- 22. Jin Z, Zhu Y, Zhang S, Xie F, Zhang M, Guo Y, et al. Diagnosis of thyroid cancer using a TI-RADS-based computer-aided diagnosis system: a multicenter retrospective study. Clinical Imaging. 2021; 80: 43-49.
- 23. Sharifi Y, Amiri Tehranizadeh A, Danay Ashgzari M, Naseri Z. TIRADS-based artificial intelligence systems for ultrasound images of thyroid nodules: protocol for a systematic review. J Ultrasound. 2024.
- 24. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Groupt. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine. 2009; 151: 264-269.
- 25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. biometrics. 1977: 159-174.
- 26. Cui Y, Fu C, Si C, Li J, Kang Y, Huang Y, et al. Analysis and Comparison of the Malignant Thyroid Nodules Not Recommended for Biopsy in ACR TIRADS and AI TIRADS With a Large Sample of Surgical Series. J Ultrasound Med. 2023; 42: 1225-1233.

- Gao Z, Chen Y, Sun P, Liu H, Lu Y. Clinical knowledge embedded method based on multi-task learning for thyroid nodule classification with ultrasound images. Phys Med Biol. 2023; 68.
- Gild ML, Chan M, Gajera J, Lurie B, Gandomkar Z, Clifton-Bligh RJ. Risk stratification of indeterminate thyroid nodules using ultrasound and machine learning algorithms. Clin Endocrinol (Oxf). 2022; 96: 646-652.
- Kunapinun A, Songsaeng D, Buathong S, Dailey MN, Keatmanee C, Ekpanyapong M. Explainable Automated TI-RADS Evaluation of Thyroid Nodules. Sensors (Basel). 2023; 23.
- 30. Lai M, Feng B, Yao J, Wang Y, Pan Q, Chen Y, et al. Value of Artificial Intelligence in Improving the Accuracy of Diagnosing TI-RADS Category 4 Nodules. Ultrasound Med Biol. 2023; 49: 2413-2421.
- Nguyen DT, Pham TD, Batchuluun G, Yoon HS, Park KR. Artificial intelligence-based thyroid nodule classification using information from spatial and frequency domains. Journal of Clinical Medicine. 2019; 8.
- 32. Song D, Yao J, Jiang Y, Shi S, Cui C, Wang L, et al. A new xAI framework with feature explainability for tumors decision-making in Ultrasound data: comparing with Grad-CAM. Comput Methods Programs Biomed. 2023; 235: 107527.
- 33. Stib MT, Pan I, Merck D, Middleton WD, Beland MD. Thyroid Nodule Malignancy Risk Stratification Using a Convolutional Neural Network. Ultrasound Q. 2020; 36: 164-172.
- Vadhiraj VV, Simpkin A, O'Connell J, Ospina NS, Maraka S, O'Keeffe DT. Ultrasound Image Classification of Thyroid Nodules Using Machine Learning Techniques. Medicina-Lithuania. 2021; 57.
- 35. Yamashita R, Kapoor T, Alam MN, Galimzianova A, Syed SA, Akdogan MU, et al. Toward Reduction in False-Positive Thyroid Nodule Biopsies with a Deep Learning-based Risk Stratification System Using US Cine-Clip Images. Radiology-Artificial Intelligence. 2022; 4.
- Ye H, Hang J, Chen X, Di X, Chen J, Ye X, et al. An intelligent platform for ultrasound diagnosis of thyroid nodules. Sci Rep. 2020; 10: 13223.
- Zhang SJ, Du HR, Jin Z, Zhu YQ, Zhang Y, Xie F, et al. A Novel Interpretable Computer-Aided Diagnosis System of Thyroid Nodules on Ultrasound Based on Clinical Experience. Ieee Access. 2020; 8: 53223-53231.
- Zhang XY, Zhang D, Han LZ, Pan YS, Wei Q, Lv WZ, et al. Predicting Malignancy of Thyroid Micronodules: Radiomics Analysis Based on Two Types of Ultrasound Elastography Images. Academic Radiology. 2023; 30: 2156-2168.
- Zhao F, Zhang H, Cheng D, Wang W, Li Y, Wang Y, et al. Predicting the risk of nodular thyroid disease in coal miners based on different machine learning models. Frontiers in Medicine. 2022; 9.
- Zhang F, Sun Y, Wu X, Meng C, Xiang M, Huang T, et al. Analysis of the Application Value of Ultrasound Imaging Diagnosis in the Clinical Staging of Thyroid Cancer. J Oncol. 2022; 2022: 8030262.
- 41. Zhuang Y, Li C, Hua Z, Chen K, Lin JL. A novel TIRADS of US classification. BioMedical Engineering Online. 2018; 17.
- 42. Zhang T, Li FX, Mu JL, Liu JT, Zhang S. Multivariate evaluation of Thyroid Imaging Reporting and Data System (TI-RADS) in diagnosis malignant thyroid nodule: application to PCA and PLS-DA analysis. International Journal of Clinical Oncology. 2017; 22: 448-454.
- Buda M, Wildman-Tobriner B, Hoang JK, Thayer D, Tessler FN, Middleton WD, et al. Management of Thyroid Nodules Seen on US Images: Deep Learning May Match Performance of Radiologists. Radiology. 2019; 292: 695-701.
- 44. Wu GG, Lv WZ, Yin R, Xu JW, Yan YJ, Chen RX, et al. Deep Learning Based on ACR TI-RADS Can Improve the Differential Diagnosis of Thyroid Nodules. Front Oncol. 2021; 11: 575166.
- Zhuo Y, Fang H, Yuan J, Gong L, Zhang Y. Fine-Needle Aspiration Biopsy Evaluation-Oriented Thyroid Carcinoma Auxiliary Diagnosis. Ultrasound Med Biol. 2023; 49: 1173-1181.
- 46. Zhou T, Hu T, Ni Z, Yao C, Xie Y, Jin H, et al. Comparative analysis of machine learning-based ultrasound radiomics in predicting malignancy of partially cystic thyroid nodules. Endocrine. 2023.

- 47. Zhang J, Wang Q, Zhao J, Yu H, Wang F, Zhang J. Automatic ultrasound diagnosis of thyroid nodules: a combination of deep learning and KWAK TI-RADS. Phys Med Biol. 2023; 68.
- Yu Z, Liu S, Liu P, Liu Y. Automatic detection and diagnosis of thyroid ultrasound images based on attention mechanism. Comput Biol Med. 2023; 155: 106468.
- 49. Yao SQ, Shen PC, Dai TW, Dai F, Wang Y, Zhang WT, et al. Human understandable thyroid ultrasound imaging AI report system-A bridge between AI and clinicians. Iscience. 2023; 26.
- 50. Xiong Z, Shi Y, Zhang Y, Duan S, Ding Y, Zheng Q, et al. Ultrasound radiomics based XGBoost model to differential diagnosis thyroid nodules and unnecessary biopsy rate: Individual application of SHapley additive exPlanations. J Clin Ultrasound. 2023.
- Zhu YC, Du H, Jiang Q, Zhang T, Huang XJ, Zhang Y, et al. Machine Learning Assisted Doppler Features for Enhancing Thyroid Cancer Diagnosis: A Multi-Cohort Study. J Ultrasound Med. 2022; 41: 1961-1974.
- 52. Zhao CK, Ren TT, Yin YF, Shi H, Wang HX, Zhou BY, et al. A Comparative Analysis of Two Machine Learning-Based Diagnostic Patterns with Thyroid Imaging Reporting and Data System for Thyroid Nodules: Diagnostic Performance and Unnecessary Biopsy Rate. Thyroid. 2021; 31: 470-481.
- 53. Xu SY, Ni XF, Zhou W, Zhan WW, Zhang H. Development and validation of a novel diagnostic tool for predicting the malignancy probability of thyroid nodules: A retrospective study based on clinical, B-mode, color doppler and elastographic ultrasonographic characteristics. Frontiers in Endocrinology. 2022; 13.
- Han D, Ibrahim N, Lu F, Zhu Y, Du H, AlZoubi A. Automatic Detection of Thyroid Nodule Characteristics From 2D Ultrasound Images. Ultrason Imaging. 2024; 46: 41-55.
- 55. Tao Y, Yu Y, Wu T, Xu X, Dai Q, Kong H, et al. Deep learning for the diagnosis of suspicious thyroid nodules based on multimodal ultrasound images. Front Oncol. 2022; 12: 1012724.
- Wei X, Gao M, Yu R, Liu Z, Gu Q, Liu X, et al. Ensemble Deep Learning Model for Multicenter Classification of Thyroid Nodules on Ultrasound Images. Med Sci Monit. 2020; 26: e926096.
- 57. Wildman-Tobriner B, Buda M, Hoang JK, Middleton WD, Thayer D, Short RG, et al. Using Artificial Intelligence to Revise ACR TI-RADS Risk Stratification of Thyroid Nodules: Diagnostic Accuracy and Utility. Radiology. 2019; 292: 112-119.
- Ataide EJG, Ponugoti N, Illanes A, Schenke S, Kreissl M, Friebe M. Thyroid Nodule Classification for Physician Decision Support Using Machine Learning-Evaluated Geometric and Morphological Features. Sensors. 2020; 20.
- Ayvaz E, Kaplan K, Kuncan F, Ayvaz E, Türkoglu H. Reducing Operation Costs of Thyroid Nodules Using Machine Learning Algorithms with Thyroid Nodules Scoring Systems. Applied Sciences-Basel. 2022; 12.
- Chen Y, Gao Z, He Y, Mai W, Li J, Zhou M, et al. An Artificial Intelligence Model Based on ACR TI-RADS Characteristics for US Diagnosis of Thyroid Nodules. Radiology. 2022; 303: 613-619.

- Cordes M, Götz TI, Coerper S, Kuwert T, Schmidkonz C. Ultrasound characteristics of follicular and parafollicular thyroid neoplasms: diagnostic performance of artificial neural network. Thyroid Research. 2023; 16.
- 62. Jin Z, Pei S, Ouyang L, Zhang L, Mo X, Chen Q, et al. Thy-Wise: An interpretable machine learning model for the evaluation of thyroid nodules. Int J Cancer. 2022; 151: 2229-2243.
- Zhang XW, Ze YY, Sang JF, Shi XB, Bi Y, Shen SM, et al. Risk factors and diagnostic prediction models for papillary thyroid carcinoma. Frontiers in Endocrinology. 2022; 13.
- Bai Z, Chang L, Yu R, Li X, Wei X, Yu M, et al. Thyroid nodules risk stratification through deep learning based on ultrasound images. Med Phys. 2020; 47: 6355-6365.
- 65. Yu H, Li JQ, Sun JL, Zheng J, Wang S, Wang GP, et al. Intelligent diagnosis algorithm for thyroid nodules based on deep learning and statistical features. Biomedical Signal Processing and Control. 2022; 78.
- Deng P, Han X, Wei X, Chang L. Automatic classification of thyroid nodules in ultrasound images using a multi-task attention network guided by clinical knowledge. Comput Biol Med. 2022; 150: 106172.
- Liu W, Lin C, Chen D, Niu L, Zhang R, Pi Z. Shape-margin knowledge augmented network for thyroid nodule segmentation and diagnosis. Comput Methods Programs Biomed. 2024; 244: 107999.
- Han X, Chang L, Song K, Cheng L, Li M, Wei X. Multitask network for thyroid nodule diagnosis based on TI-RADS. Med Phys. 2022; 49: 5064-5080.
- 69. Ha EJ, Lee JH, Lee DH, Na DG, Kim JH. Development of a machine learning-based fine-grained risk stratification system for thyroid nodules using predefined clinicoradiological features. Eur Radiol. 2023; 33: 3211-3221.
- Long J, Shelhamer E, Darrell T, editors. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition; 2015.
- Gu Z, Cheng J, Fu H, Zhou K, Hao H, Zhao Y, et al. Ce-net: Context encoder network for 2d medical image segmentation. IEEE transactions on medical imaging. 2019; 38: 2281-2292.
- 72. Choi YJ, Baek JH, Park HS, Shim WH, Kim TY, Shong YK, et al. A computeraided diagnosis system using artificial intelligence for the diagnosis and characterization of thyroid nodules on ultrasound: initial clinical assessment. Thyroid. 2017; 27: 546-552.
- Wang H, Yang II Y, Peng III B, Chen IV Q, editors. A thyroid nodule classification method based on TI-RADS. Ninth International Conference on Digital Image Processing (ICDIP 2017); 2017: SPIE.
- 74. Pedraza L, Vargas C, Narváez F, Durán O, Muñoz E, Romero E, editors. An open access thyroid ultrasound image database. 10th International symposium on medical information processing and analysis; 2015: SPIE.
- Zhou TH, Hu T, Ni ZK, Yao C, Xie YY, Jin HM, et al. Comparative analysis of machine learning-based ultrasound radiomics in predicting malignancy of partially cystic thyroid nodules. Endocrine. 2023.
- Sharifi Y, Shafiei S, Ashgzari MD, Zakavi SR, Eslami S. Thyroid Ultrasound-Image Dataset. Challenges of Trustable AI and Added-Value on Health: IOS Press; 2022: 397-402.