International Journal of Innovative Technologies in Social Science

e-ISSN: 2544-9435

Operating Publisher
SciFormat Publishing Inc.
ISNI: 0000 0005 1449 8214

2734 17 Avenue SW,
Calgary, Alberta, T3E0A7,
Canada
+15878858911
editorial-office@sciformat.ca

ARTICLE TITLE	ARTIFICIAL INTELLIGENCE IN ULTRASONOGRAPHIC DIAGNOSIS OF THYROID NODULES: ENHANCING RISK STRATIFICATION AND CLINICAL DECISON-MAKING
DOI	https://doi.org/10.31435/ijitss.1(49).2026.4860
RECEIVED	09 December 2025
ACCEPTED	25 January 2026
PUBLISHED	02 February 2026
LICENSE	The article is licensed under a **Creative Commons Attribution 4.0 International License.**

ARTIFICIAL INTELLIGENCE IN ULTRASONOGRAPHIC DIAGNOSIS OF THYROID NODULES: ENHANCING RISK STRATIFICATION AND CLINICAL DECISON-MAKING

Michał Kociński *(Corresponding Author, Email: michal.kocinski1999@gmail.com)*
Jan Biziel University Hospital No. 2 in Bydgoszcz, Bydgoszcz, Poland
ORCID ID: 0009-0007-7651-7929

Filip Matusiak
Jan Biziel University Hospital No. 2 in Bydgoszcz, Bydgoszcz, Poland
ORCID ID: 0009-0002-0538-6443

Klaudia Brzoza
Jan Biziel University Hospital No. 2 in Bydgoszcz, Bydgoszcz, Poland
ORCID ID: 0009-0006-0950-4514

Patryk Iglewski
Wojewódzki Szpital Zespolony im. L. Rydygiera w Toruniu, Toruń, Poland
ORCID ID: 0009-0004-6611-2168

Michał Pietrasz
Wojewódzki Szpital Zespolony im. L. Rydygiera w Toruniu, Toruń, Poland
ORCID ID: 0009-0000-8148-7487

Anna Komarczewska
Wojewódzki Szpital Zespolony im. L. Rydygiera w Toruniu, Toruń, Poland
ORCID ID: 0009-0006-7378-2607

ABSTRACT

Thyroid nodules are common clinical findings, increasingly detected due to the widespread use of high-resolution ultrasound imaging. While the majority of these nodules are benign, a minority may be malignant, necessitating accurate and efficient risk stratification. Traditional ultrasonographic evaluation relies heavily on the operator's expertise and subjective interpretation, which introduces diagnostic variability.

This narrative review explores the evolving role of artificial intelligence in the ultrasonographic diagnosis of thyroid nodules. The principal objective of this review is to critically evaluate the diagnostic performance, clinical utility, and integration potential of artificial intelligence (AI)-based methodologies—including machine learning (ML) and deep learning (DL)—in the ultrasonographic assessment of thyroid nodules. Particular attention is devoted to the enhancement of existing risk stratification frameworks, and to identifying barriers to implementation in routine clinical. The review evaluates AI-integrated diagnostic systems in relation to existing classification frameworks, such as the thyroid imaging reporting and data system, and highlights innovations in elastography, 3D imaging, and automated segmentation. Evidence suggests that AI can enhance diagnostic accuracy, reduce interobserver variability, and improve the standardization of thyroid nodule assessment. Some algorithms demonstrate performance comparable to that of experienced clinicians, particularly in differentiating benign from suspicious nodules.

Despite promising results, limitations such as model generalizability, the need for large annotated datasets, and clinical validation remain challenges. The findings support the integration of artificial intelligence as a complementary tool to assist healthcare professionals in making more objective, consistent, and timely decisions regarding the evaluation and management of thyroid nodules

Methodology: A comprehensive narrative review of peer-reviewed literature was undertaken, encompassing both classical and AI-augmented ultrasonographic techniques, with a specific focus on diagnostic criteria, algorithmic accuracy, and classification consistency across TI-RADS variants (ACR-TIRADS, EU-TIRADS, K-TIRADS). Additionally, the role of emerging modalities such as ultrasound elastography was examined in the context of evaluating cytologically indeterminate nodules. Literature published between 2009 and 2025 was examined to assess how machine learning and deep learning algorithms contribute to image interpretation, classification, and malignancy prediction.

Abbreviated Description of The State Of Knowledge: Thyroid nodules are detected in up to 60% of the general adult population via ultrasonography. Although the malignancy rate remains relatively low (~5%), the clinical imperative is the accurate differentiation of malignant from benign lesions. Risk stratification relies on the assessment of sonographic features, including echogenicity, shape, margins, calcifications, and vascularity. Several standardized scoring systems—most notably TI-RADS—are employed to systematize malignancy risk and guide indications for fine-needle aspiration biopsy (FNAB). Despite its utility, ultrasonography remains inherently operator-dependent and subject to interpretive variability. AI-powered diagnostic systems have demonstrated promising potential in mitigating interobserver discrepancies, augmenting risk classification fidelity, and improving diagnostic throughput. Adjunctive techniques such as elastography provide additional biomechanical data, although limitations in methodological standardization currently preclude widespread adoption.

KEYWORDS

Thyroid Nodules, High-Resolution Ultrasonography, Artificial Intelligence, Machine Learning, Deep Learning, TI-RADS, EU-TIRADS, Diagnostic Standardization, Fine-Needle Aspiration, Elastography, Thyroid Malignancy

CITATION

Michał Kociński, Filip Matusiak, Klaudia Brzoza, Patryk Iglewski, Michał Pietrasz, Anna Komarczewska. (2026) Artificial Intelligence in Ultrasonographic Diagnosis of Thyroid Nodules: Enhancing Risk Stratification and Clinical Decision-Making. *International Journal of Innovative Technologies in Social Science*. 1(49). doi: 10.31435/ijitss.1(49).2026.4860

1. Introduction

Thyroid nodules represent a frequent clinical finding, increasingly identified through high-resolution ultrasound imaging. While the majority of these lesions are benign, the diagnostic process must reliably distinguish malignant nodules to guide appropriate management and avoid unnecessary interventions [1,2]. Despite the widespread use of ultrasonography as the primary imaging modality for initial assessment, its diagnostic accuracy remains limited by interobserver variability and dependence on the clinician's experience [3,4].

In recent years, artificial intelligence (AI) has emerged as a transformative tool in medical imaging, offering advanced capabilities in automated pattern recognition, feature extraction, and diagnostic classification [3,5]. Machine learning (ML) and deep learning (DL) algorithms, in particular, have demonstrated promising results in supporting risk stratification systems such as TI-RADS and improving diagnostic consistency [1,5,6]. These tools may help refine clinical decisions, reduce the frequency of unnecessary fine-needle aspiration biopsies, and assist in identifying nodules with malignant potential more effectively [6,7].

Against this backdrop, this study aims to evaluate the clinical applicability and diagnostic performance of AI in the ultrasonographic assessment of thyroid nodules, compare it with expert-based evaluations, and outline the challenges and future directions for its integration into clinical practice.

2. Overview of thyroid nodules as a common clinical issue

Thyroid nodules (TN) are abnormal masses within the thyroid gland, often discovered incidentally during routine physical examinations or imaging tests [8,31]. Their prevalence in the general population is remarkably high, with high-resolution ultrasonography detecting TN in up to 60% of individuals. The likelihood of developing TN increases with certain risk factors, including female sex, advancing age, and higher body mass index (BMI) [2,9].

TN can manifest in various forms, ranging from benign colloid nodules and thyroid cysts to more complex conditions such as multinodular goiter, hyperfunctioning nodules, and thyroid cancer [8]. 90% of all thyroid cancers are comprised by differentiated thyroid cancer (including papillary and follicular cancer) [10].

Although most TN are asymptomatic, some patients may notice a palpable lump in the neck. In certain cases, larger nodules can exert pressure on adjacent structures, leading to compressive symptoms such as dysphagia, dyspnea, or hoarseness. While the majority of TN are non-functioning, some nodules may produce excessive thyroid hormones, resulting in symptoms of hyperthyroidism [8].

Approximately 95% of TN are benign [8]. However, because around 5% of cases are malignant, the primary clinical concern when evaluating a thyroid nodule is the exclusion of thyroid cancer [2,10]. Several factors increase the risk of malignancy, including exposure to ionizing radiation, rapid nodule growth, hoarseness, and a family history of thyroid cancer or other cancers. Additionally, patients with Graves' disease demonstrate a significantly increased prevalence of papillary thyroid cancer, with reported rates ranging between 33% and 42% [7].

The evaluation of a newly diagnosed TN is essential to determine whether it is benign or malignant and, in cases of symptomatic benign nodules, to assess the need for intervention to alleviate associated local symptoms [3]. A comprehensive assessment involves multiple steps, including a detailed clinical history and physical examination, laboratory tests to measure thyroid function and detect thyroid antibodies (thyroid peroxidase and thyroid stimulating hormone receptor antibody), and imaging studies like thyroid ultrasound. In some cases, additional imaging, such as computed tomography of the neck, may be required. When malignancy is suspected, fine-needle aspiration or biopsy is recommended [8].

To aid in clinical decision making, various scoring systems have been developed, including the Thyroid Imaging Reporting and Data System (TIRADS). This classification system assesses nodules to estimate the probability of malignancy [8].

3. Ultrasonography (USG) in thyroid nodule evaluation.

Ultrasonography (US) plays a pivotal role in the evaluation of thyroid nodules, outperforming other imaging modalities in terms of sensitivity for detecting focal lesions [2,13]. The widespread use of neck ultrasonography has led to a marked increase in the detection of thyroid nodules [1,2,19] and thyroid carcinomas [14], although a substantial proportion of these lesions are benign and require only periodic follow-up [1,2,19]. Notably, thyroid nodules are most often discovered incidentally during imaging performed for unrelated reasons [1].

The longitudinal monitoring of thyroid nodules necessitates a detailed analysis of specific ultrasonographic features, including the nodule's composition, echogenicity, margins, shape, and the presence of echogenic or hyperechoic foci. The size of the nodule remains a controversial parameter in determining the need for further diagnostic procedures [11]. Nodule morphology, as assessed by ultrasonography, may serve as an indication for fine-needle aspiration biopsy (FNAB) [1,14].

It is important to acknowledge that different ultrasonographic classification systems are employed worldwide to stratify the risk of malignancy in thyroid nodules. The most commonly used include the American College of Radiology Thyroid Imaging Reporting and Data System (ACR-TIRADS), the European Thyroid Association system (EU-TIRADS), and the Korean Thyroid Imaging Reporting and Data System (K-TIRADS). These systems differ in their scoring algorithms and the weight assigned to specific morphological features, which may influence clinical decision-making regarding further diagnostic and therapeutic approaches [12]. Nevertheless, all three classification systems demonstrate high diagnostic accuracy and are suitable for use in routine clinical practice, emphasizing that ACR-TIRADS is distinguished by high specificity values, EU-TIRADS and K-TIRADS are distinguished by high values. [21].

Despite its diagnostic value, ultrasonography has inherent limitations. As a highly operator-dependent modality [22], the interpretation of key features of thyroid nodule — such as hypoechogenicity, microcalcifications, irregular margins, and vascularity—may vary significantly between observers. While these features are associated with a higher risk of malignancy, they are not sufficiently specific or reliable to establish a definitive diagnosis. Therefore, additional diagnostic investigations are often warranted such as FNA or core needle biopsy (CNB) by experienced operators [21]. In cases of suspected malignancy, further imaging modalities such as radionuclide scintigraphy or positron emission tomography-computed tomography (PET-CT), as well as cytological assessment via FNAB, are indispensable. Ultrasonography alone should not be considered a sufficient basis for clinical diagnosis [13].

Furthermore, ultrasound elastography has emerged as a valuable adjunct in the assessment of thyroid nodules. This technique evaluates tissue stiffness using either internal or external mechanical stimuli or shear wave propagation. Elastography may aid in identifying malignant lesions prior to FNAB and is particularly useful in cases with indeterminate cytological findings. However, the clinical utility of elastography is currently limited by a lack of standardization, underscoring the need for further validation and harmonization of this method [23].

4. TIRADS

TI-RADS (Thyroid Imaging Reporting and Data System) is an ultrasound-based classification system designed to evaluate focal thyroid lesions and assess their risk of malignancy. Each category within the system corresponds to a specific level of risk and provides recommendations for further management, including the need for fine-needle aspiration (FNA). Originally introduced in 2009 [15], the TI-RADS system has since undergone multiple revisions and improvements based on clinical evidence and practical experience. The most recent and widely adopted version in Europe is EU-TIRADS, developed by the European Thyroid Association [16]. TI-RADS relies on the assessment of sonographic features such as echogenicity, shape, margins, presence of microcalcifications, and nodule size. Unlike point-based systems, EU-TIRADS uses a pattern recognition approach to classify nodules into five risk categories.

Numerous studies have shown that thyroid ultrasound is the imaging modality of choice for identifying nodules with suspicious features. The EU-TIRADS classification demonstrates high sensitivity and substantial specificity in predicting malignancy [18,19].

TI-RADS is a simple, clear, and practical tool that greatly facilitates clinical decision-making and improves workflow in daily medical practice. An additional benefit is the availability of online EU-TIRADS calculators [17], which help automate the classification process and speed up diagnosis.

Results and Methods

1. Introduction to Artificial Intelligence and Machine Learning in Medical Imaging

Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has increasingly become a transformative force in medical imaging. This shift is fueled by a confluence of factors, including the rapid growth of imaging data, advancements in computational power, and the development of sophisticated algorithmic architectures capable of directly interpreting complex data patterns. Although AI is not a new concept, recent breakthroughs in neural networks and the availability of annotated medical datasets have ushered in a new era of performance and applicability across clinical imaging domains [20,25].

Medical imaging lends itself particularly well to AI integration due to the repetitive and structured nature of image interpretation tasks. Traditionally, image analysis relied on manually engineered features and rule-based systems, which often lacked robustness and generalizability. The emergence of convolutional neural networks (CNNs) has marked a paradigm shift—these networks learn hierarchical features directly from raw image inputs and have demonstrated superior performance in classification, segmentation, object detection, and image enhancement [20,21].

Recent advances have expanded the AI toolbox beyond CNNs. Reinforcement learning (RL), for instance, has been explored in medical image analysis as a dynamic framework in which models learn optimal decision-making policies through interactions with their environment. This has shown promise in image registration, lesion localization, and multi-step image interpretation workflows [22].

The clinical utility of AI in imaging is increasingly supported by empirical evidence. Several comparative studies have demonstrated that deep learning models can achieve diagnostic performance on par with human experts in tasks such as detecting malignancies in radiographs, segmenting organs, and classifying skin lesions [23]. Nevertheless, external validation remains inconsistent, and model performance may vary considerably across institutions due to differences in imaging protocols, equipment, and patient demographics.

One of the most critical limitations of current AI systems is their generalizability. Models trained on data from a single institution often underperform when applied to external datasets—a problem exacerbated by a lack of standardized imaging protocols and representative training data. Moreover, the opacity of many AI models (the so-called "black-box" problem) continues to challenge clinical acceptance and trust [24].

Ethical considerations also come to the forefront. AI models can unintentionally replicate and amplify biases present in training data, potentially leading to unequal diagnostic outcomes across populations. This has led to growing interest in fairness-aware machine learning and the development of regulatory frameworks to ensure equitable and transparent AI integration in clinical environments [23,24].

Furthermore, medical images differ significantly from natural photographs—they are often three-dimensional, multi-sequence, and encoded in DICOM formats that include metadata and context-specific information. Integrating these modalities into AI pipelines while maintaining patient privacy and clinical relevance adds another layer of complexity [21].

Despite these challenges, AI is unlikely to replace radiologists. Rather, its role is increasingly envisioned as augmentative-helping to streamline workflows, reduce inter-observer variability, prioritize cases based on urgency, and provide decision support in routine and complex cases alike. Importantly, ongoing interdisciplinary collaboration among clinicians, computer scientists, ethicists, and regulators is essential to fully realize the potential of AI in a safe, effective, and human-centered manner [23,25].

2. The Role of Artificial Intelligence in the Ultrasonographic Diagnosis of Thyroid Nodules

Thyroid nodules are a very common phenomenon—detectable in up to approximately 65% of the population, especially due to the widespread use of ultrasonography in diagnostics. Although most of them are benign, about 10% can be malignant, which necessitates prompt and accurate diagnosis. In recent decades, the incidence of thyroid cancer has increased, likely due to improved diagnostic methods and higher detection rates. In clinical practice, the evaluation of thyroid nodules is primarily based on ultrasonographic examination; however, this method is subjective and depends on the operator's experience. Hence, there is a need for more objective and repeatable approaches, and artificial intelligence (AI) is increasingly becoming such a solution [26].

3. Traditional Machine Learning and Deep Learning in Ultrasonographic Diagnosis

Two main AI approaches are used in diagnosing thyroid nodules: traditional machine learning (ML) and deep learning (DL).

Machine learning involves manual feature extraction—such as shape, texture, and echo intensity—from identified regions of interest (ROI). Based on these features, classifiers like Support Vector Machines (SVM) can distinguish between benign and malignant nodules. A drawback of this method is the need for expert knowledge to manually define and select relevant features [26].

Deep learning, particularly Convolutional Neural Networks (CNNs), can learn features independently without manual definition by analyzing raw ultrasound images directly. Thanks to transfer learning—using models pre-trained on large, general datasets such as ImageNet—models can be adapted to medical images even with limited training data. DL-based systems achieve diagnostic performance comparable to, or even exceeding, experienced radiologists using TIRADS (Thyroid Imaging Reporting and Data System) classifications [26].

Studies have shown that combining both methods (a so-called fusion model)—integrating results from traditional ML and deep transfer learning—significantly improves diagnostic accuracy. In a study involving 630 patients, the fusion of an SVM model and a deep learning model based on the InceptionV3 architecture achieved an AUC (area under the curve) of 0.783, a significant improvement over either model used alone [26].

4. Advanced Dynamic Image Analysis and 3D Visualization

Most traditional methods rely on analyzing static images (single 2D slices). However, thyroid nodules are three-dimensional structures, and analyzing dynamic ultrasound videos and reconstructing them in 3D allows for more accurate assessment.

One example is the TNVis (Thyroid Nodules Visualization) system, which uses advanced deep learning algorithms to segment nodules in ultrasound videos and then builds 3D models. This system was tested on over 5,000 cases and achieved high segmentation accuracy (Dice Similarity Coefficient – 0.90). Importantly, TNVis improved radiologists' diagnostic performance, increasing AUC from 0.66 (without AI) to 0.79 (with full AI support). Junior physicians in particular benefited, improving their accuracy by over 15% [27].

This system enables better assessment of the shape and size of nodules, which is crucial for malignancy risk evaluation and clinical planning.

5. Multimodal Approach: Elastography, TIRADS, and Radiomics Supported by Interpretable Machine Learning

To improve diagnostic quality, AI integrates various imaging methods and scoring systems. Elastography is an ultrasound technique that measures tissue stiffness—malignant lesions are often stiffer than benign ones.

By combining elastography, TIRADS classification, and advanced radiomic feature extraction, a model called ELTIRADS was developed, which integrates this data using an SVM algorithm. In a study of 181 patients, ELTIRADS achieved very high diagnostic performance: 92% accuracy, 89% sensitivity, and 94% specificity—exceeding diagnosis based solely on physician assessment [28].

Additionally, interpretable methods such as Shapley Additive Explanations (SHAP) were employed to show the contribution of individual features to the model's final decision. This helps physicians better understand AI recommendations and increases trust in the system [28].

6. Clinical Validation and Real-World Application

An important step is confirming that AI models work not only in research conditions but also in daily clinical practice. A logistic regression model incorporating clinical data (e.g., TSH level, presence of autoimmune thyroiditis) and detailed ultrasound features (stiffness, echogenicity, microcalcifications, shape) was validated on an independent cohort of 455 patients. It demonstrated strong diagnostic performance (high c-statistic), good calibration, and practical clinical utility—reducing the number of unnecessary biopsies and surgeries [29].

7. Modern Neural Network Architectures and Data Integration

Contemporary deep learning techniques such as U-Net and Mask R-CNN are used to precisely segment nodules in ultrasound images with accuracy exceeding 93%. Transfer learning enables the use of these networks even with limited medical data.

Moreover, AI systems are evolving toward integrating clinical, imaging, and laboratory data, enabling not only malignancy risk diagnosis but also disease progression forecasting and individualized treatment planning for thyroid cancer [30].

Artificial intelligence significantly enhances the ultrasonographic diagnosis of thyroid nodules by combining traditional machine learning, advanced deep learning, dynamic and 3D imaging analysis, elastography, and interpretable models. This leads to more objective and effective cancer risk assessment, reduces unnecessary biopsies, and improves the overall quality of patient care.

8. Comparison of the Effectiveness of AI and Radiologists in Thyroid Ultrasound Diagnosis

A 2021 study showed that radiologists from the British Thyroid Association and the American College of Radiology TIRADS demonstrate similar diagnostic effectiveness compared to an artificial intelligence model called Artificial Intelligence TIRADS. Both groups were characterized by high sensitivity but low specificity in lesion assessment. However, radiologists had a higher rate of unnecessarily performed fine-needle biopsies [31].

Technological advancement, however, is relentless. A 2024 study showed that artificial intelligence systems may achieve significantly better performance than humans. A comparative analysis of diagnostic accuracy was conducted using a dataset of 600 ultrasound images of the thyroid, verified by cytological examination. The diagnostic performance of experienced radiologists, third-year radiology residents, and an AI model was compared.

Experienced radiologists achieved a sensitivity of 40.0%, specificity of 80.0%, accuracy of 63.3%, positive predictive value of 58.8%, and negative predictive value of 65.1%. Residents achieved a sensitivity of 64.0%, specificity of 77.1%, accuracy of 71.7%, positive predictive value of 66.7%, and negative predictive value of 75.0%.

The Artificial Intelligence Model demonstrated a sensitivity of 80.0%, specificity of 71.4%, accuracy of 75.0%, positive predictive value of 66.7%, and negative predictive value of 83.3% [32]. In another study, the diagnostic accuracy of a group of 80 radiologists with varying experience levels was compared with an AI model trained on 2,069 thyroid ultrasound images representing 655 different thyroid nodules. A test set of 100 thyroid nodule cases was used. The artificial intelligence system achieved 73% diagnostic accuracy, while the average accuracy among radiologists was 62.9% [33].

Discussion

The integration of artificial intelligence into the ultrasonographic diagnosis of thyroid nodules represents a significant advancement in precision medicine, particularly in addressing the limitations of operator-dependent and subjective image interpretation. As demonstrated in multiple studies, AI—especially deep learning models such as convolutional neural networks—can offer diagnostic accuracy that rivals or surpasses that of experienced radiologists. This is particularly noteworthy given the high prevalence of thyroid nodules and the clinical importance of differentiating benign from malignant lesions with minimal invasiveness.

One of the most compelling advantages of AI lies in its ability to process and learn from large volumes of imaging data, thereby reducing variability and standardizing assessments across institutions and practitioners. Traditional machine learning methods, while valuable, often depend on manual feature engineering and expert-defined inputs. Deep learning, by contrast, enables end-to-end learning from raw image data, allowing the automatic extraction of complex and high-dimensional features. The application of transfer learning has further facilitated the adoption of these models in medical imaging, even in the context of limited domain-specific data.

Furthermore, the evolution of multimodal approaches—integrating elastography, radiomic features, and established scoring systems like TIRADS—has enhanced diagnostic specificity and sensitivity. Models such as ELTIRADS, which incorporate interpretable machine learning techniques like SHAP, not only improve performance metrics but also address a major concern in clinical AI: explainability. This transparency is essential for fostering clinician trust and ensuring accountability in diagnostic decision-making.

The development of systems such as TNVis illustrates how dynamic image analysis and 3D reconstruction can further augment traditional ultrasound imaging. By providing a more comprehensive visualization of nodular structures, these tools assist in better malignancy risk stratification and may be particularly beneficial in training and supporting less experienced clinicians. Notably, the observed improvements in diagnostic performance among junior physicians suggest a potential role for AI in medical education and decision support.

However, despite these promising developments, several challenges must be acknowledged. First, the generalizability of AI models remains a concern. Diagnostic performance often declines when models are applied to data from external institutions with different imaging protocols, equipment, or patient demographics. This underscores the need for diverse and representative training datasets, as well as robust external validation. Ethical considerations are also paramount: the risk of algorithmic bias, data privacy issues, and the potential for overreliance on automated systems necessitate cautious and transparent deployment of AI in clinical settings.

In addition, while comparative studies indicate that AI can outperform or at least match radiologist performance in thyroid nodule evaluation, the goal is not to replace human expertise but to complement it. The most effective applications of AI are likely to be those that augment clinical workflows—prioritizing cases, flagging suspicious features, and providing second opinions—thereby improving efficiency and reducing diagnostic errors without undermining human judgment.

In summary, AI holds substantial promise for improving the ultrasonographic assessment of thyroid nodules by increasing diagnostic accuracy, standardization, and efficiency. Future research should focus on addressing issues of generalizability, ensuring algorithmic fairness, and exploring the integration of AI into broader clinical decision-making frameworks. Continued interdisciplinary collaboration will be vital in translating these technological advances into sustainable clinical practice.

Conclusions

Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has emerged as a transformative force in medical imaging, driven by the exponential growth of imaging data, advances in computational capabilities, and increasingly sophisticated algorithmic models. In the context of thyroid ultrasonography—a critical tool in evaluating the malignancy risk of thyroid nodules—AI offers the potential to overcome key limitations of conventional diagnostic approaches, which are often subjective and operator-dependent.

Traditional ML methods rely on manual feature extraction from ultrasound images, requiring domain expertise to select relevant characteristics such as shape, texture, and echogenicity. In contrast, DL models, especially convolutional neural networks (CNNs), automatically learn complex features from raw data and have demonstrated performance levels comparable to, and sometimes exceeding, those of experienced radiologists using systems like TIRADS. Fusion models combining ML and DL further enhance diagnostic accuracy.

Recent advancements include the use of AI in dynamic image analysis and 3D visualization, exemplified by systems such as TNVis, which significantly improve both the precision of nodule segmentation and diagnostic outcomes, particularly among less experienced clinicians. AI integration with elastography and radiomics has yielded models like ELTIRADS, which exhibit high sensitivity, specificity, and interpretability through tools such as SHAP, facilitating clinical trust and transparency.

Real-world validation studies confirm AI's ability to reduce unnecessary biopsies and optimize patient management by integrating clinical, imaging, and laboratory data. Furthermore, modern architectures such as U-Net and Mask R-CNN enhance segmentation accuracy and enable individualized treatment planning.

Comparative studies suggest that AI models now rival and occasionally surpass human experts in diagnostic performance. While experienced radiologists demonstrate high specificity, AI systems often achieve superior sensitivity and overall accuracy, particularly in standardized datasets. Nonetheless, challenges remain regarding generalizability, data heterogeneity, ethical implications, and the need for transparent and fair algorithms.

In conclusion, AI holds considerable promise as an augmentative tool in the ultrasonographic assessment of thyroid nodules, offering enhanced objectivity, efficiency, and diagnostic precision. Ongoing interdisciplinary collaboration and robust clinical validation will be essential to ensure its responsible and equitable implementation in healthcare practice.

Authors contribution:

Conceptualization: Michał Kociński, Filip Matusiak, Klaudia Brzoza, Patryk Iglewski, Michał Pietrasz, Anna Komarczewska

Methodology: Anna Komarczewska, Patryk Iglewski, Michał Pietrasz

Software: not applicable

Verification: Klaudia Brzoza,, Anna Komarczewska, Patryk Iglewski

Formal analysis: Michał Kociński, Filip Matusiak, Michał Pietrasz

Research: Michał Kociński, Klaudia Brzoza, Patryk Iglewski,, Anna Komarczewska, Filip Matusiak, Michał Pietrasz

Data curation: Michał Kociński, Michał Pietrasz, Filip Matusiak

Resources: Patryk Iglewski, Anna Komarczewska, Michał Pietrasz

Writing-rough preparation: Filip Matusiak, Patryk Iglewski Michał Kociński

Writing-review and editing: Michał Kociński, Filip Matusiak, Klaudia Brzoza, Patryk Iglewski, Michał Pietrasz, Anna Komarczewska

Supervision: Klaudia Brzoza, Michał Pietrasz, Anna Komarczewska

Project administration: Filip Matusiak, Patryk Iglewski, Michał Pietrasz, Anna Komarczewska, Michał Kociński, Klaudia Brzoza

All authors have read and agreed with published version of the manuscript.

Funding statement: This research receive no external founding

Conflict of Interests Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

REFERENCES

1. Rago, T., & Vitti, P. (2022). Risk Stratification of Thyroid Nodules: From Ultrasound Features to TIRADS. *Cancers*, *14*(3), 717. https://doi.org/10.3390/cancers14030717
2. Grani, G., Sponziello, M., Pecce, V., Ramundo, V., & Durante, C. (2020). Contemporary Thyroid Nodule Evaluation and Management. *The Journal of clinical endocrinology and metabolism*, *105*(9), 2869–2883. https://doi.org/10.1210/clinem/dgaa322
3. Bernet, V. J., & Chindris, A. M. (2021). Update on the Evaluation of Thyroid Nodules. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, *62*(Suppl 2), 13S–19S. https://doi.org/10.2967/jnumed.120.246025
4. Alyami, J., Almutairi, F. F., Aldoassary, S., Albeshry, A., Almontashri, A., Abounassif, M., & Alamri, M. (2022). Interobserver variability in ultrasound assessment of thyroid nodules. *Medicine*, *101*(41), e31106. https://doi.org/10.1097/MD.0000000000031106
5. Moraes, P. H. M., Sigrist, R., Takahashi, M. S., Schelini, M., & Chammas, M. C. (2019). Ultrasound elastography in the evaluation of thyroid nodules: evolution of a promising diagnostic tool for predicting the risk of malignancy. *Radiologia brasileira*, *52*(4), 247–253. https://doi.org/10.1590/0100-3984.2018.0084
6. Tobcu, E., Karavaş, E., Yılmaz, G. T., & Topçu, B. (2025). Comparison of K-TIRADS, EU-TIRADS and ACR-TIRADS Guidelines for Malignancy Risk Determination of Thyroid Nodules. *Diagnostics (Basel, Switzerland)*, *15*(8), 1015. https://doi.org/10.3390/diagnostics15081015
7. Kant, R., Davis, A., & Verma, V. (2020). Thyroid Nodules: Advances in Evaluation and Management. *American family physician*, *102*(5), 298–304.
8. Hamill, C., Ellis, P., & Johnston, P. C. (2022). Ultrasound for the assessment of thyroid nodules: an overview for non-radiologists. *British journal of hospital medicine (London, England : 2005)*, *83*(7), 1–7. https://doi.org/10.12968/hmed.2022.0071
9. Uppal, N., Collins, R., & James, B. (2023). Thyroid nodules: Global, economic, and personal burdens. *Frontiers in endocrinology*, *14*, 1113977.https://doi.org/10.3389/fendo.2023.1113977
10. Haugen, B. R., Alexander, E. K., Bible, K. C., Doherty, G. M., Mandel, S. J., Nikiforov, Y. E., Pacini, F., Randolph, G. W., Sawka, A. M., Schlumberger, M., Schuff, K. G., Sherman, S. I., Sosa, J. A., Steward, D. L., Tuttle, R. M., & Wartofsky, L. (2016). 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 : official journal of the American Thyroid Association*, *26*(1), 1–133. https://doi.org/10.1089/thy.2015.0020
11. Burgos, N., Ospina, N. S., & Sipos, J. A. (2022). The Future of Thyroid Nodule Risk Stratification. *Endocrinology and metabolism clinics of North America*, *51*(2), 305–321. https://doi.org/10.1016/j.ecl.2021.12.002

12. Piticchio, T., Russ, G., Radzina, M., Frasca, F., Durante, C., & Trimboli, P. (2024). Head-to-head comparison of American, European, and Asian TIRADSs in thyroid nodule assessment: systematic review and meta-analysis. *European thyroid journal*, *13*(2), e230242. https://doi.org/10.1530/ETJ-23-0242
13. Harbuwono, D. S., Soewondo, P., Yunir, E., Soebardi, S., Darmowidjojo, B., Purnamasari, D., Tarigan, T. J. E., Wisnu, W., Tahapary, D. L., Kurniawan, F., Yulian, E. D., Lisnawati, L., Stephanie, A., Makes, B., Zulkarnaien, B., Suroyo, I., Siswoyo, A. D., Gondhowiardjo, S., Kodrat, H., Hermani, B., … Subekti, I. (2019). Diagnostic Approach for Thyroid Nodules. *Acta medica Indonesiana*, *51*(2), 189–193.
14. Shin, J. H., Baek, J. H., Chung, J., Ha, E. J., Kim, J. H., Lee, Y. H., Lim, H. K., Moon, W. J., Na, D. G., Park, J. S., Choi, Y. J., Hahn, S. Y., Jeon, S. J., Jung, S. L., Kim, D. W., Kim, E. K., Kwak, J. Y., Lee, C. Y., Lee, H. J., Lee, J. H., … Korean Society of Thyroid Radiology (KSThR) and Korean Society of Radiology (2016). Ultrasonography Diagnosis and Imaging-Based Management of Thyroid Nodules: Revised Korean Society of Thyroid Radiology Consensus Statement and Recommendations. *Korean journal of radiology*, *17*(3), 370–395. https://doi.org/10.3348/kjr.2016.17.3.370
15. Horvath, E., Majlis, S., Rossi, R., Franco, C., Niedmann, J. P., Castro, A., & Dominguez, M. (2009). An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. *The Journal of clinical endocrinology and metabolism*, *94*(5), 1748–1751. https://doi.org/10.1210/jc.2008-1724
16. Durante, C., Hegedüs, L., Czarniecka, A., Paschke, R., Russ, G., Schmitt, F., Soares, P., Solymosi, T., & Papini, E. (2023). 2023 European Thyroid Association Clinical Practice Guidelines for thyroid nodule management. *European thyroid journal*, *12*(5), e230067. https://doi.org/10.1530/ETJ-23-0067
17. *Ti-rads calculator – calculates ti-rads score*. https://tiradscalculator.com/
18. Bukasa-Kakamba, J., Bayauli, P., Sabbah, N., Bidingija, J., Atoot, A., Mbunga, B., Nkodila, A., Atoot, A., Bangolo, A. I., & M'Buyamba-Kabangu, J. R. (2022). Ultrasound performance using the EU-TIRADS score in the diagnosis of thyroid cancer in Congolese hospitals. *Scientific reports*, *12*(1), 18442. https://doi.org/10.1038/s41598-022-22954-y
19. Jin, Z., Pei, S., Shen, H., Ouyang, L., Zhang, L., Mo, X., Chen, Q., You, J., Zhang, S., & Zhang, B. (2023). Comparative Study of C-TIRADS, ACR-TIRADS, and EU-TIRADS for Diagnosis and Management of Thyroid Nodules. *Academic radiology*, *30*(10), 2181–2191. https://doi.org/10.1016/j.acra.2023.04.013
20. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005
21. Prevedello, L. M., Halabi, S. S., Shih, G., Wu, C. C., Kohli, M. D., Chokshi, F. H., Erickson, B. J., Kalpathy-Cramer, J., Andriole, K. P., & Flanders, A. E. (2019). Challenges Related to Artificial Intelligence Research in Medical Imaging and the Importance of Image Analysis Competitions. *Radiology. Artificial intelligence*, *1*(1), e180031. https://doi.org/10.1148/ryai.2019180031
22. Hu, M., Zhang, J., Matkovic, L., Liu, T., & Yang, X. (2023). Reinforcement learning in medical image analysis: Concepts, applications, challenges, and future directions. *Journal of applied clinical medical physics*, *24*(2), e13898. https://doi.org/10.1002/acm2.13898
23. Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., Mahendiran, T., Moraes, G., Shamdas, M., Kern, C., Ledsam, J. R., Schmid, M. K., Balaskas, K., Topol, E. J., Bachmann, L. M., Keane, P. A., & Denniston, A. K. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *The Lancet. Digital health*, *1*(6), e271–e297. https://doi.org/10.1016/S2589-7500(19)30123-2
24. Koçak, B., Ponsiglione, A., Stanzione, A., Bluethgen, C., Santinha, J., Ugga, L., Huisman, M., Klontzas, M. E., Cannella, R., & Cuocolo, R. (2025). Bias in artificial intelligence for medical imaging: fundamentals, detection, avoidance, mitigation, challenges, ethics, and prospects. *Diagnostic and interventional radiology (Ankara, Turkey)*, *31*(2), 75–88. https://doi.org/10.4274/dir.2024.242854
25. Avanzo, M., Stancanello, J., Pirrone, G., Drigo, A., & Retico, A. (2024). The Evolution of Artificial Intelligence in Medical Imaging: From Computer Science to Machine and Deep Learning. *Cancers*, *16*(21), 3702. https://doi.org/10.3390/cancers16213702
26. Xu, Y., Xu, M., Geng, Z., Liu, J., & Meng, B. (2025). Thyroid nodule classification in ultrasound imaging using deep transfer learning. *BMC cancer*, *25*(1), 544. https://doi.org/10.1186/s12885-025-13917-3
27. Zhou, Y., Chen, C., Yao, J., Yu, J., Feng, B., Sui, L., Yan, Y., Chen, X., Liu, Y., Zhang, X., Wang, H., Pan, Q., Zou, W., Zhang, Q., Lin, L., Xu, C., Yuan, S., He, Q., Ding, X., Liang, P., … Xu, D. (2025). A deep learning based ultrasound diagnostic tool driven by 3D visualization of thyroid nodules. *NPJ digital medicine*, *8*(1), 126. https://doi.org/10.1038/s41746-025-01455-y
28. Barzegar Golmoghani, Erfan & Mohebi, Mobin & Gohari, Zahra & Aram, Sadaf & Mohammadzadeh, Ali & Firouznia, Sina & Shakiba, Madjid & Naghibi, Hamed & Moradian, Sadegh & Ahmadi, Maryam & Almasi, Kazhal & Issaiy, Mahbod & Anjomrooz, Mehran & Tavangar, Seyed Mohammad & Javadi, Sheida & Bitarafan Rajabi, Ahmad & Davoodi, Mohammad & Sharifian, Hashem & Mohammadzadeh, Maryam. (2025). ELTIRADS framework for thyroid nodule classification integrating elastography, TIRADS, and radiomics with interpretable machine learning. Scientific Reports. 15. 10.1038/s41598-025-93226-8.

29. Fernández Alba, J. J., Carral, F., Ayala Ortega, C., Santotoribio, J. D., Lara, M. C., & González Macías, C. (2025). External Validation of a Predictive Model for Thyroid Cancer Risk with Decision Curve Analysis. *Diagnostics (Basel, Switzerland)*, *15*(6), 686. https://doi.org/10.3390/diagnostics15060686
30. Yang, L., Wang, X., Zhang, S., Cao, K., & Yang, J. (2025). Research progress on artificial intelligence technology-assisted diagnosis of thyroid diseases. *Frontiers in oncology*, *15*, 1536039. https://doi.org/10.3389/fonc.2025.1536039
31. Watkins, L., O'Neill, G., Young, D., & McArthur, C. (2021). Comparison of British Thyroid Association, American College of Radiology TIRADS and Artificial Intelligence TIRADS with histological correlation: diagnostic performance for predicting thyroid malignancy and unnecessary fine needle aspiration rate. *The British journal of radiology*, *94*(1123), 20201444. https://doi.org/10.1259/bjr.20201444
32. Namsena, P., Songsaeng, D., Keatmanee, C., Klabwong, S., Kunapinun, A., Soodchuen, S., Tarathipayakul, T., Tanasoontrarat, W., Ekpanyapong, M., & Dailey, M. N. (2024). Diagnostic performance of artificial intelligence in interpreting thyroid nodules on ultrasound images: a multicenter retrospective study. *Quantitative imaging in medicine and surgery*, *14*(5), 3676–3694. https://doi.org/10.21037/qims-23-1650
33. Liu, Y., Feng, Y., Qian, L., Wang, Z., & Hu, X. (2023). Deep learning diagnostic performance and visual insights in differentiating benign and malignant thyroid nodules on ultrasound images. *Experimental biology and medicine (Maywood, N.J.)*, *248*(24), 2538–2546. https://doi.org/10.1177/15353702231220664