Artificial Intelligence for Thyroid Nodule Characterization: Where Are We Standing?

Salvatore Sorrenti [1,†], Vincenzo Dolcetti [2,†], Maija Radzina [3,4], Maria Irene Bellini [1,*], Fabrizio Frezza [5,6], Khushboo Munir [5], Giorgio Grani [7], Cosimo Durante [7], Vito D'Andrea [1], Emanuele David [7], Pietro Giorgio Calò [8], Eleonora Lori [1] and Vito Cantisani [2]

1   Department of Surgical Sciences, "Sapienza" University of Rome, 00161 Rome, Italy;
    salvatore.sorrenti@uniroma1.it (S.S.); vito.dandrea@uniroma1.it (V.D.); eleonora.lori@uniroma1.it (E.L.)
2   Department of Radiological, Anatomo-Pathological Sciences, "Sapienza" University of Rome,
    00161 Rome, Italy; vincenzodolcetti@gmail.com (V.D.); vito.cantisani@uniroma1.it (V.C.)
3   Radiology Research Laboratory, Riga Stradins University, LV-1007 Riga, Latvia; mradzina@gmail.com
4   Medical Faculty, University of Latvia, Diagnostic Radiology Institute, Paula Stradina Clinical University
    Hospital, LV-1007 Riga, Latvia
5   Department of Information Engineering, Electronics and Telecommunications, "Sapienza" University of Rome,
    00184 Rome, Italy; fabrizio.frezza@uniroma1.it (F.F.); khushboo.muniruniroma1@gmail.com (K.M.)
6   Consorzio Nazionale Interuniversitario per le Telecomunicazioni (CNIT), Viale G.P. Usberti 181/A Sede
    Scientifica di Ingegneria-Palazzina 3, 43124 Parma, Italy
7   Department of Translational and Precision Medicine, "Sapienza" University of Rome, 00161 Rome, Italy;
    giorgio.grani@uniroma1.it (G.G.); cosimo.durante@uniroma1.it (C.D.); emanuele.david@uniroma1.it (E.D.)
8   Department of Surgical Sciences, "Policlinico Universitario Duilio Casula", University of Cagliari,
    09042 Monserrato, Italy; pgcalo@unica.it
*   Correspondence: mariairene.bellini@uniroma1.it
†   These authors contributed equally to this work.

Simple Summary: In the present review, an up-to-date summary of the state of the art of artificial intelligence (AI) implementation for thyroid nodule characterization and cancer is provided. The opinion on the real effectiveness of AI systems remains controversial. Taking into consideration the largest and most scientifically valid studies, it is possible to state that AI provides results that are comparable or inferior to expert ultrasound specialists and radiologists. Promising data approve AI as a support tool and simultaneously highlight the need for a radiologist supervisory framework for AI provided results. Therefore, current solutions might be more suitable for educational purposes.

Abstract: Machine learning (ML) is an interdisciplinary sector in the subset of artificial intelligence (AI) that creates systems to set up logical connections using algorithms, and thus offers predictions for complex data analysis. In the present review, an up-to-date summary of the current state of the art regarding ML and AI implementation for thyroid nodule ultrasound characterization and cancer is provided, highlighting controversies over AI application as well as possible benefits of ML, such as, for example, training purposes. There is evidence that AI increases diagnostic accuracy and significantly limits inter-observer variability by using standardized mathematical algorithms. It could also be of aid in practice settings with limited sub-specialty expertise, offering a second opinion by means of radiomics and computer-assisted diagnosis. The introduction of AI represents a revolutionary event in thyroid nodule evaluation, but key issues for further implementation include integration with radiologist expertise, impact on workflow and efficiency, and performance monitoring.

Keywords: artificial intelligence; machine learning; thyroid cancer

1. Introduction

For thyroid nodule management, the current diagnostic goal is early identification of the malignant thyroid nodules: although the incidence of the disease is high (incidence rate

of 3.4/100,000 in men and 11.5/100,000 in women [1]), more than half of newly diagnosed thyroid cancers have a low risk of persistence or recurrence [2,3]. It is therefore necessary to develop a diagnostic tool that improves interobserver agreement in the risk stratification of thyroid nodules to provide an objective assessment of utility for the clinical and surgical management phases that follow [4], given that even molecular biology is not specific and does not accurately predict prognosis after surgery [5,6].

In the last two decades, medical imaging has grown exponentially, shifting from the traditional use of images for visual interpretation to their conversion to quantitative features that can be analyzed to extrapolate data and thus improve clinical decision-making. This approach is usually called "Radiomics" [7,8]. Radiomics takes advantage from extraction algorithms to derive several quantitative features from radiological images. Several recent works underline how these data may be used by machine learning (ML) systems.

ML is an interdisciplinary sector in the subset of artificial intelligence (AI) dealing with the creation of systems that set up logical connections via algorithms to make predictions on data systems [9], Figure 1. The most interesting application of ML in the medical field is the discernment of patterns based on the examination and analysis of extensive datasets coming from various sources (clinical databases, laboratory results, and imaging data) [10,11]. In particular, ML techniques are divided into supervised and unsupervised learning methods. Supervised ML uses dataset inputs linked to dataset (labeled) outputs to identify a function between the two, while unsupervised ML uses non-labeled input datasets to identify and separate subsets with similar characteristics [12].
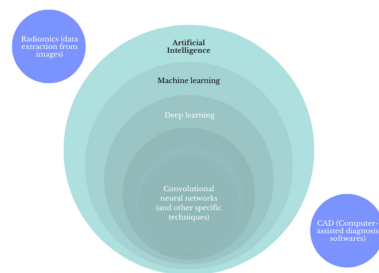

Figure 1. Schematic definition of artificial intelligence, machine learning, deep learning, and convolutional neural networks.

Deep learning (DL) is subset of ML approaches that uses neural networks arranged in layers to extract higher level features from input data and automatically learn their discriminative features, which allows approximation of non-linear relationships with excellent performance.

These technologies may be finally transferred to software used directly by clinicians: Computer Aided Diagnosis (CAD). Such software can be stand-alone or integrated in sonographic equipment and help in the detection and evaluation of thyroid nodules, one of the most common endocrine diseases, with incidental finding on ultrasound (US) examination, especially in patients over 65 years of age [13].

2. Materials and Methods

The study only considered articles published in the last decade (2012–2022), since most of the literature concerning AI application in radiology has undergone extensive development only recently. Among these, only large retrospective and prospective studies, systematic reviews, and meta-analyses were selected, as overall, they have greater statistical significance. The research was carried out by interrogating the PubMed and Google Scholar online databases using the Mesh terms "thyroid nodule and artificial intelligence", with the MESH terms present in the titles or abstracts. Only human studies were selected. The search identified 166 studies from January 2012 to April 2022; of these, 63 were further considered. After a full text read, the final studies included in the review were 30 in number; they are all listed below in Table 1 [14–42].

Table 1. Machine learning approaches for the evaluation of thyroid nodule sonographic images.

Reference	Approach	Source Data	Method Details	Performance
Zhu, et al., 2021 [14]	Brief Efficient Thyroid Network (BETNET; a CSS model)	gray-scale US images of 592 patients with 600 TNs (internal dataset) 187 patients with 200 TNs (external validation dataset)	CNN approach with 24 layers: 13 convolution layers, 5 pooling layers, 3 fully connected layers with dropouts in between	AUC 0.970, 95% CI: 0.958–0.980 in the independent validation cohort; similar to two highly skilled radiologists (0.940 and 0.953)
Peng, et al. 2021 [15]	Deep-learning AI model (ThyNet)	18,049 US images of 8339 patients (training set) 4305 images of 2775 patients (total test set)	combined architecture of three networks: ResNet, ResNeXt, and DenseNet	ThyNet AUC (0.922; 95% CI 0.910–0.934] higher than that of the radiologists (0.839; CI 0.834–0.844]; $p < 0.0001$)
Bai, et al., 2021 [16]	RS-Net evaluation AI model	13,984 thyroid US images	CNN approach in which GoogLeNet is used as the backbone network.	Accuracy, sensitivity, specificity, PPV, and NPV were 88.0%, 98.1%, 79.1%, 80.5%, and 97.9%, comparable to that of a senior radiologist
Yoon, et al., 2021 [17]	Texture analysis; least absolute shrinkage and selection operator (LASSO) logistic regression model including clinical variables	155 US images of indeterminate thyroid nodules in 154 patients.	Texture extraction using MATLAB 2019b.; the LASSO model was used to choose the most useful predictive features. Univariable and multivariable logistic regression analyses were performed to build malignancy prediction models.	Integrated model AUC 0.839 vs. 0.583 (clinical variables only).
Liu, et al., 2021 [18]	information fusion-based joint convolutional neural network (IF-JCNN)	163 pairs of US images and raw radiofrequency signals of thyroid nodules	IF-JCNN contains two branched CNNs for deep feature extraction: one for US images (14 convolutional layers and 3 fully connected layers) and the other one for RF signals (12 convolutional layers and 3 fully connected layers)	The information carried by raw radiofrequency signals and ultrasound images for thyroid nodules is complementary IF-JCNN (both images and RF signals): AUC 0.956 (95% CI 0.926–0.987)
Gomes Ataide, et al., 2020 [19]	Feature extraction and Random Forest classifier	99 original US images	Feature extraction using MATLAB 2018b; Random Forest classifier (400 Decision Trees; Criterion: Entropy, with Bootstrap)	RFC accuracy 99.3%, sensitivity 99.4%, specificity 99.2%
Ye, et al., 2020 [20]	Deep convolution neural network (VGG-16)	US images of 1601 nodules (training set) and test data including 209 nodules (test set)	CNN approach based on VGG-19 (16 layers with learnable weights, 13 convolutions and 3 fully connected layers)	AUC 0.9157, comparable to the experienced radiologist (0.8879; $p > 0.1$)
Wei, et al., 2020 [21]	Ensemble deep learning model (EDLC-TN)	25,509 thyroid US images	CNN model based on DenseNet and adopted as a multistep cascade pathway for an ensemble learning model with voting system.	AUC 0.941 (0.936–0.946)
Zhou, et al., 2020 [23]	CNN-based transfer learning method named DLRT (deep-learning radiomics of thyroid)	US images of 1750 thyroid nodules (from 1734 patients)	CNN-based architecture with transfer learning strategy, with 4 hidden layers (3 transferred and a fine-tuned layer) and a fully connected layer	AUC in the external cohort 0.97 (0.95–0.99). Both a senior and a junior US radiologist had lower sensitivity and specificity than DLRT.

Table 1. *Cont.*

Reference	Approach	Source Data	Method Details	Performance
Nguyen, et al., 2020 [24]	Combination of multiple CNN models (ResNet-based and InceptionNet-based)	450 US thyroid nodule images (from 298 patients)	Combination of ResNet50-based (50 layers) and Inception-based (4 layers) networks followed by global average pooling, batch normalization, dropout, and dense layer	Accuracy: 92.05%
Wang, et al., 2020 [25]	Three CNN networks (feature extraction network; attention-based feature aggregation network; classification network)	7803 US thyroid nodule images from 1046 examinations	CNN approach based on Inception-Resnet-v2 (164 layers)	Method AUC 0.9006 Both the accuracy and sensitivity are significantly higher than sonographers.
Thomas, et al., 2020 [26]	AIBx, AI model to risk stratify thyroid nodules	2025 US images of 482 thyroid nodules (internal dataset) and 103 nodules (external dataset)	CNN approach based on ResNet 34 (34 layers)	Negative predictive value (NPV), sensitivity, specificity, positive predictive value (PPV), and accuracy of the image similarity model were greater than other cancer risk stratification systems.
Galimzianova, et al., 2020 [27]	Feature extraction and regularized logistic regression model	92 US images of 92 biopsy-confirmed thyroid nodules	Feature extraction (219 for each nodule) and elastic net regression analysis	Method AUC 0.828 (95% CI, 0.715–0.942), greater than or comparable to that of the expert classifiers
Nguyen, et al., 2019 [28]	AI-Based Thyroid Nodule Classification Using Information from Spatial and Frequency Domains	ultrasound thyroid images of 237 patients (training dataset) and 61 patients (test dataset).	CNN models (Resnet18, Resnet34, and Resnet50 were compared)	AI system with spatial domain based on deep learning, and frequency domain based on Fast Fourier transform (FFT) outperforms the state-of-the-art methods (especially CAD systems)
Buda, et al., 2019 [29]	CNN	1377 US images of thyroid nodules in 1230 patients (training dataset) and 99 nodules (internal test dataset)	Custom CNN (six blocks with 3×3 convolutional filters, followed by Rectified Linear Unit activation function and max pooling layer with 2×2 kernels).	Method AUC: 0.87 [CI 0.76, 0.95] Three ACR-TIRADS readers 0.91
Koh, et al., 2020 [30]	Two individual CNNs compared with experienced radiologist	15,375 US images of thyroid nodules (training set), 634 (internal test), 1181 (external test set).	Four CNNs including two individual CNNs, ResNet50 (50 layers) and InceptionResNetV2 (164 layers), and two classification ensembles, AlexNet-GoogLeNet-SqueezeNet ensemble and AlexNet-GoogLeNetSqueezeNet-InceptionResNetv2 ensemble	CNNs AUC similar to experienced radiologist AUC (0.87)
Wang, et al., 2019 [31]	CNN compared with experienced radiologist	351 US images with nodules and 213 images without nodules of 276 patients	CNN system in which the Resnet v2-50 (50 layers) network and YOLOv2 are integrated	CAD AUC 0.902 significantly higher than radiologist AUC 0.859 ($p = 0.0434$)

Table 1. *Cont.*

Reference	Approach	Source Data	Method Details	Performance
CAD systems				
Sun, et al., 2020 [22]	Fused features combing the CNN-based features (VGG F-based features) with hand-crafted features	1037 US images of thyroid nodules (internal dataset) and 550 images (test dataset)	A support vector machine (SVM) is used for classification and fused features which combined the deep features extracted by a CNN with hand-crafted features, such as the histogram of oriented gradient (HOG), local binary patterns (LBP), and scale invariant feature transform (SIFT)	AUC of attending radiology lower than system (0.819 vs. 0.881, $p = 0.0003$)
Han, et al., 2021 [32]	S-Detect for Thyroid	US images of 454 thyroid nodules from 372 consecutive patients	S-Detect for Thyroid is an AI-based CAD software integrated in US equipment (Samsung Medison Co., Seoul, South Korea)	The sensitivities of the CAD system did not differ significantly from those of the radiologist (all $p > 0.05$); the specificities and accuracies were significantly lower than those of the radiologist (all $p < 0.001$).
Zhang, et al., 2020 [33]	AI-SONIC; Demetics Medical Technology Co., Zhejiang, China	US images of 365 thyroid nodules	AI-SONIC is a CAD based on deep learning (cascade CNN of two different CNN architectures (one with 15 convolutional layers/2 pooling layers for segmentation, and the other with 4 convolutional layers/4 pooling layers for detection), developed by Demetics Medical Technology Co., China	AUC CAD 0.788 vs. senior radiologist 0.906, $p < 0.001$. The use of CAD system improved the diagnostic sensitivities of both the senior and the junior radiologists
Fresilli, et al., 2020 [4]	S-Detect for Thyroid compared with an expert radiologist, a senior resident and a medical student evaluation	US images of 107 thyroid nodules	S-Detect for Thyroid is an AI-based CAD software integrated in US equipment (Samsung Medison Co., Seoul, South Korea)	The CAD system and the expert achieved similar values of a sensitivity and specificity (about 70%–87.5%). The specificity achieved by the student was significantly lower (76.25%).
Jin, et al., 2020 [34]	CAD system based on a modified, CNN-based TIRADS, evaluated by	US images of 789 thyroid nodules from 695 patients	CAD system based on the ACR TI-RADS automatic scoring using a CNN (no details provided).	AUC CAD 0.87 AUC Junior radiologist 0.73 (Junion + CAD): 0.83 AUC Senior radiologist 0.91
Xia, et al., 2019 [35]	S-Detect for Thyroid	US images of 180 thyroid nodules in 171 consecutive patients	S-Detect for Thyroid is an AI-based CAD software integrated in US equipment (Samsung Medison Co., Seoul, South Korea)	AUC CADs 0.659 (0.577–0.740) AUC radiologist 0.823 (0.758–0.887)
Jin, et al., 2019 [36]	AmCad; AmCad BioMed, Taipei City, Taiwan	33 images from 33 patients read by 81 radiologists	Commercial standalone CAD software: AmCad (version: Shanghai Sixth People's Hospital; AmCad BioMed, Taipei City, Taiwan)	CAD AUC 0.985 (0.881–1.00) 177 contestants AUC 0.659 (0.645–0.673) ($p < 0.01$)

Table 1. *Cont.*

Reference	Approach	Source Data	Method Details	Performance
Kim, et al., 2019 [37]	S-Detect for Thyroid 1 and 2	US images of 218 thyroid nodules from 106 consecutive patients	S-Detect for Thyroid is an AI-based CAD software integrated in US equipment (Samsung Medison Co., Seoul, South Korea)	AUC: radiologist 0.905 (95% CI, 0.859–0.941) S-Detect 1–assisted radiologist 0.865 (0.812–0.907) S-Detect 1 0.814 (0.756–0.863) S-Detect 2-assisted radiologist 0.802 (0.743–0.853) S-Detect 2 0.748 (0.685–0.804)
Chi, et al., 2017 [38]	CAD system for thyroid nodule	Database 1 includes 428 images in total while database 2 includes 164 images in total	CAD based on fine tuning of GoogLeNet CNN (22 convolutional layers including 9 inception modules)	CAD AUC 0.9920 Experienced radiologist AUC 0.9135
Zhao, et al., 2019 [39]	CAD system for thyroid nodule systematic review and meta-analysis	Meta-analysis of 5 studies with 723 thyroid nodules from 536 patients	4 studies with S-Detect; 1 study with internal CAD based on CNN.	CAD AUC 0.90 (95% CI 0.87–0.92) Experienced radiologist AUC 0.96 (95% CI 0.94–0.97)
AI-modified TIRADS				
Watkins, et al., 2021 [40]	AI-TIRADS	US images of 218 nodules from 212 patients	The AI-TIRADS is an optimization of ACR TIRADS generated by "genetic algorithms", a subgroup of AI methods that focus on algorithms inspired by "natural selection".	Sensitivity 93.44% Specificity 45.71% BTA, ACR-TIRADS, and AI-TIRADS have comparable diagnostic performance
Wang, et al., 2020 [41]	Google AutoML for automated nodule identification and risk stratification	US images of 252 nodules from 249 patients.	Google AutoML algorithm (AutoML Vision; Google LLC), with cloud computing and transfer learning	Accuracy of 68.7 ± 7.4% of AI-integrated TIRADS
Wildman-Tobriner, et al., 2010 [42]	AI-TIRADS	US images of 1425 biopsy-proven thyroid nodules from 1264 consecutive patients (training set); 100 nodules (test set)	The AI-TIRADS is an optimization of ACR TIRADS generated by "genetic algorithms", a subgroup of AI methods that focus on algorithms inspired by "natural selection".	ACR TI-RADS AUC 0.91 AI TI-RADS AUC 0.93 (with slight improvement of specificity and ease of use)

Abbreviations: ACR: American College of Radiology; AI: artificial intelligence; AIBx: AI model to risk stratify thyroid nodules; AUC: area under the curve; AutoML: Auto machine learning; BETNET: brief efficient thyroid network; CAD: computer-aided diagnosis; CI: confidence interval; CNN: convolution neural network; CSS: cascading style sheets; DLRT: deep-learning radiomics of thyroid; EDLC-TN: ensemble deep-learning classification model for thyroid nodules; FFT: Fast Fourier transform; IF-JCNN: information fusion-based joint convolutional neural network; LASSO: Least Absolute Shrinkage and Selection Operator; NPV: negative predictive value; PPV: positive predictive value; RF: radiofrequency; RFC: Random Forest classifier; RS-NET: regression–segmentation network; US: ultrasound; VGG: Visual Geometry Group.

3. Results

3.1. Radiomics

Medical radiomics employs high-throughput automated extraction algorithms to obtain a large number of quantitative characteristics from image datasets and is able to identify measurable information that clinical evaluation alone cannot detect [12,43].

Two of the first radiomics approaches in thyroid nodule characterization were texture analysis and US echo-intensity evaluation [44]. The latter is affected by several factors, such as gain, dynamics, operator dependency, and probe variability, as well as by the US equipment performance. The diagnostic value of echo-intensity obtained by direct measurement is limited; however, the echo intensity of the nodule and surrounding tissues increases or decreases simultaneously when these factors alternate [45]. Therefore, the echo intensity of the thyroid nodule can be indirectly quantified by measuring the grayscale ratio of the nodule to the surrounding thyroid tissues, which is more objective than the subjective assessment [44–46]. In a pivotal single-center study, it was demonstrated that the ratio was significantly lower in malignant nodules compared to benign ones [46], while the ratio of the nodule to the strap muscle was influenced by gender and less clinically discriminant. The inter-rater agreement was fair (k = 0.40) for hypo-echogenicity, whereas it was substantial for the ratio (k = 0.74), confirming the reduction in variability. This approach was subsequently replicated by other groups, showing that, as suggested, the ratio may distinguish anechoic and markedly hypoechoic nodules [47], and if it is applied to different nodule sizes [48], software can differentiate between benign and malignant nodules [49], even in different settings [45]. One of the most significant examples is the multicenter study conducted by Liang et al., in which a radiomic score was compared with a score based on the ACR TI-RADS criteria (which take into account, in addition to the difference in echogenicity, characteristics such as composition, shape, margin, and echogenic foci), showing a close correlation between the latter and the assessment carried out by the AI [50]. Radiomics approaches using grayscale histogram and other more complex image analyses were furthermore proved to predict BRAF mutational status [51], lateral lymph node metastasis [52], and a disease-free survival term.

3.2. Deep Learning and Machine Learning and TIRADS Systems

Deep learning (DL) is one ML method that relies on networks of computational units (i.e., neural units arranged in layers that gradually extract higher-level features from input data and automatically learn discriminative features from data) that allow approximation of complex non-linear relationships with outstanding performance. DL can achieve diagnosis automation, avoiding human intervention. In medical applications, DL algorithms are implemented for detection and characterization of tissue lesions as well as for the analysis of disease progression [12].

AI has already been widely used in thyroid imaging [11,53]. Several AI and ML approaches were implemented for the classification of thyroid nodules and the early detection of cancers, including modifications to the American College of Radiology Thyroid Imaging Reporting and Data System (TIRADS) systems that may be manually applied. Furthermore, a convolutional neural-network-based CAD program may help in predicting the BRAFV600E genetic mutation [54–56].

Use of the ML approach may also identify nodules with high-risk mutations on molecular testing [57]. Another important advantage of AI systems is the possibility to obtain more systematized results, which could reduce inter-observer variability and tend to standardize the results obtained through the application of different TIRADS classification systems, whose major limit to date is represented by highly variable predictive capacity, high heterogeneity in grading, and the absence of reliable data in small nodules (<10 mm) [3,58,59] (Figures 2 and 3). A recent TIRADS model showed higher accuracy than a model based on training according to the nodule status, i.e., benign and malignant; additionally, the specificity of the above-mentioned model was higher than that of both experienced and junior radiologists [60].

Comparisons between different imaging modalities are represented in Figures 2 and 3, where a DL-based software confirms the suspect based on B-mode US imaging.
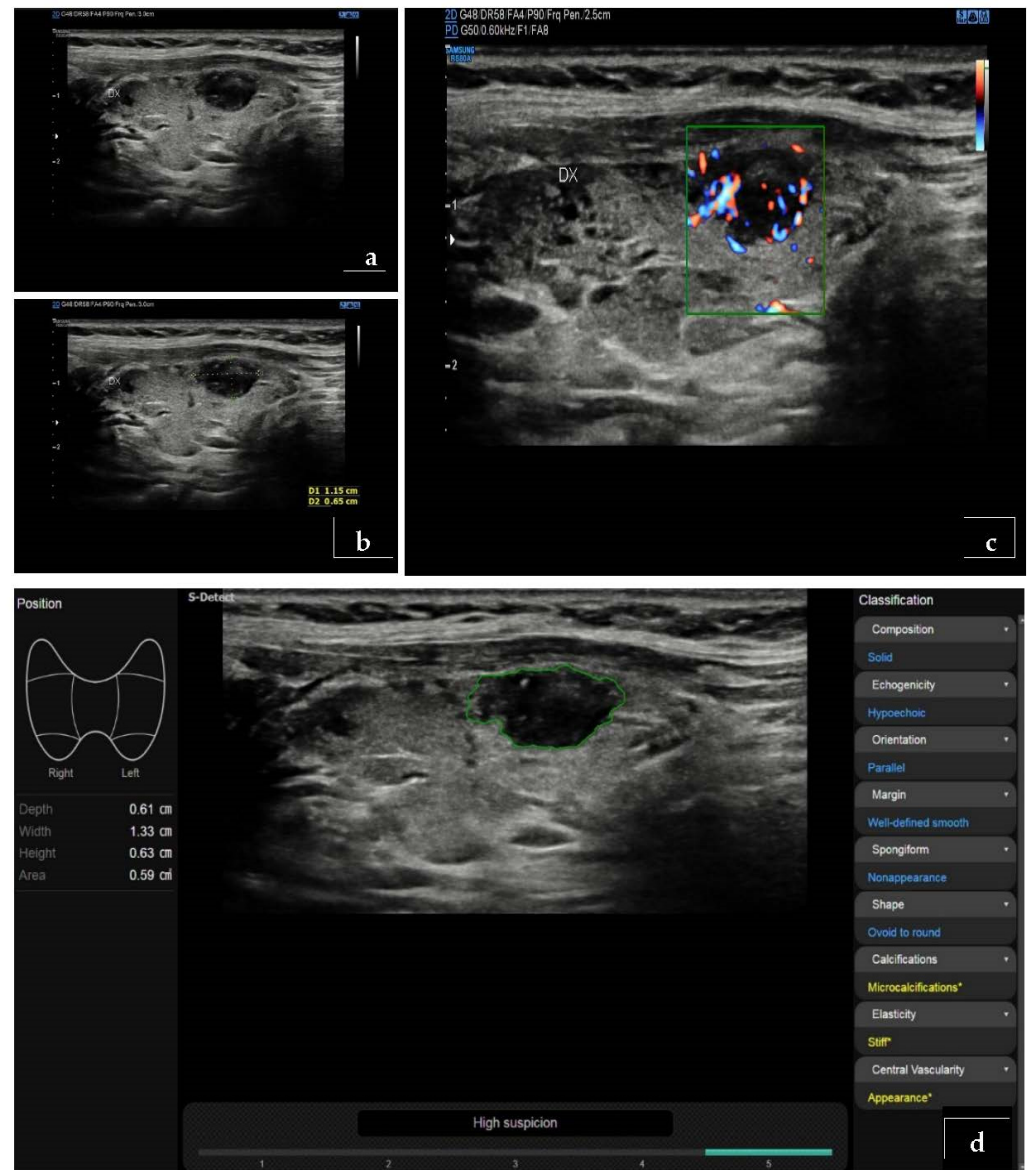

Figure 2. Histologically proved papillary carcinoma. (**a**,**b**) At B-mode, the nodule appears markedly hypoechoic. (**c**) At ColorDoppler-US evaluation, it shows type III vascular pattern (intra- and perinodular). (**d**) The evaluation with S-detect software (Samsung Medison, Co., Ltd., Seoul, Korea) confirms the high degree of suspicion (K-TIRADS 5).

3.3. Computer-Assisted Diagnosis (CAD)

These approaches may produce new knowledge by identifying new patterns and features to be applied in a more traditional way and generating computer-assisted diagnosis (CAD) systems; i.e., software able to analyze data through the application of machine-learning principles to aid clinicians for a "second opinion" provision. AI-based thyroid CADs may further improve diagnostic performance and reliability, reaching an accuracy similar to that obtained by an expert radiologist [10,11], with potential implication in training of less-experienced operators and reduction of intra- and inter-observer variability [11].
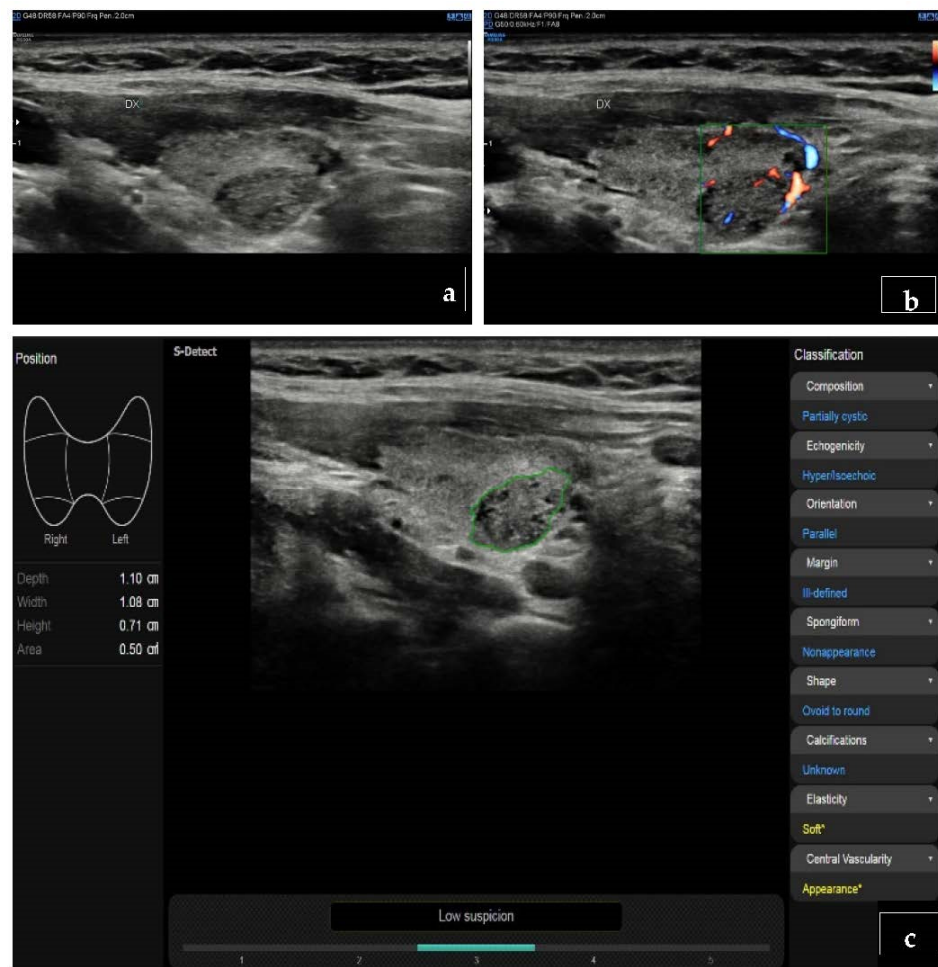
Figure 3. Histologically proved follicular hyperplasia. (**a**) The nodule appears iso-hypoechoic at B-mode. (**b**) At ColorDoppler-US evaluation, it shows type III vascular pattern (intra- and perinodular) (**c**) At S-detect software, the nodule is classified as low degree of suspicion (K-TIRADS 3).

CAD-systems are already available as commercial applications or where embedded in US equipment. A recent meta-analysis [61] confirmed that their performance in evaluating malignant thyroid nodules is comparable to radiologists. Specifically, the sensitivity was reported to be like that of experienced radiologists, while specificity and diagnostic odds ratio were reduced [39]. While these systems did not outperform experienced specialists, they are able to guide the training of less-skilled examiners, thus reducing variability when clinician's judgements show significant disagreement. However, it is difficult to eliminate all possible sources of inter-observer variability: it is in fact possible that radiologists with different degrees of experience select images with more or less relevant characteristics of suspicion. The homogeneity of the image segmentation process also plays a fundamental role in reducing the impact of selection bias. The segmentation process in fact involves a manual selection of the area of interest (which should correspond to the nodule), but in this phase it is possible that portions of the slide that contain non-informative areas are selected, compromising the training process of the AI system. To try to solve the problem, some studies have adopted a two-step fully automated classification system, specifically trained both to autonomously select the area of interest and to predict the final pathology of the specific selected area [62]

Furthermore, the models generated by images obtained from different machines may not be universally generalizable, which can determine limits in the sampling phases and in the standardization of software. This therefore requires an accurate evaluation and

selection phase prior to the adoption of an AI system in any case [11]. Table 2 summarizes main advantages and disadvantages of artificial intelligence over conventional imaging.

Table 2. Advantages and disadvantages of artificial intelligence over conventional imaging.

Main Advantages of AI	Main Disadvantages of AI
It is based on models, for the interpretation of thyroid nodules, that are able to match the performance characteristics of radiologists and pathologists	Too little experience at the moment; prospective multicenter trials on a wide population will be needed to improve the utility of artificial intelligence for the interpretation of thyroid nodules
Usable software for thyroid nodule risk stratification are already commercially available	

4. Discussion

The TIRADS system was developed to improve the diagnostic accuracy of conventional US in thyroid nodule characterization [63]. However, its clinical use is still very limited and diverse; in particular, there are various types of TIRADS, and their application is very subjective; therefore, it is significantly affected by inter-observer variability [64].

AI could increase US accuracy and significantly limit inter-observer variability by using standardized mathematical algorithms. In the world of DL, many authors are focusing on convolutional neural networks (CNNs), introduced by LeCun [65,66]. Before their diagnostic accuracy can be assessed, CNNs are trained by subjecting them to specific algorithm-segmented US images of thyroid nodules with known histological diagnosis; at the end of the learning phase the CNNs are able to analyze the captures of thyroid nodules and to suggest a risk stratification of these nodules in correlation to a specific TI-RADS level [16]. Most of the existing literature evaluates the diagnostic accuracy of various types of properly trained convolutional neural networks by comparing them to those of radiologists with variable degrees of experience. All the evaluated studies showed significant high overall diagnostic accuracy of CNNs, above 90%, which does not differ much from that of expert radiologists. In particular, most of the studies demonstrate a comparable diagnostic accuracy, such as Watkins et al., Bai et al., Ye et al., Koh et al., and Fresilli et al. [4,16,20,30,40]. Approximately the same number of studies demonstrate a higher diagnostic accuracy of AI systems compared to that of expert radiologists (e.g., Sun et al., Peng et al., and Zhou et al.) [15,22,23], or vice versa, a superiority of diagnostic accuracy by expert radiologists compared to that of AI systems (e.g., Zhang et al. and Han et al.) [32,33]. Despite controversial results, the meta-analysis conducted by Zhao et al. suggests that the sensitivity of the CAD system is like that of experienced radiologists, but the CAD system has lower specificity and diagnostic odds ratio than experienced radiologists [39].

On the other hand, almost all the studies included in this review show that CNNs obtain a better result than junior radiologists with less than 5 years of experience in US evaluation of thyroid nodules [4,23,34,40], especially with regards to specificity [60]. These studies therefore agree in suggesting that CAD systems may be an effective support tool to increase the diagnostic efficacy of thyroid nodule evaluations by less-experienced radiologists [25]. Furthermore, some studies, such as the one by Zhao et al., show that the diagnostic accuracy of senior radiologists assisted by CAD systems is higher than that of radiologists alone and CAD systems alone [39].

It is therefore not yet clear from the literature analysis which of the specific AI systems has the best diagnostic accuracy. Wang et al. compare the effectiveness of only few CNNs [25], while most studies analyze specific systems individually, showing high specificity—especially if they are based on TIRADS system algorithm—rather than differentiation among benign and malignant nodules with surgical histopathological reference [60]. In absolute terms, the CAD system used by Zhou et al., a CNN-based transfer learning method named DLRT (deep-learning radiomics of thyroid), appears to be one of those with

greater diagnostic accuracy (AUC 0.97) [23], although this type of comparison between AI systems has no real statistical significance as they were analyzed on retrospective datasets.

In addition, a variety of AI technologies have been evaluated on thyroid cytology specimens. Unfortunately, no application has been demonstrated to be robust enough for clinical use in FNAB result analysis, an issue which is related to the multi-layered, multi-dimensional, complex interpretation process and the lack of standardized algorithms [66,67]. However, Ippolito et al. [68] show collaborative data between cytology and US; they integrated microscopic pathology characteristics, clinical data, and imaging features into a combined algorithm to triage indeterminate and follicular lesions into high- or low-risk categories using a CNN framework that demonstrated a sensitivity of 85.7% and low specificity of 58.8%. As an element of evidence that emerged from the present review, key issues in AI implementation include integration with radiologist interpretation, impact on workflow and efficiency, and performance monitoring. This can be translated into an automated structured report for integration into a radiology report. Sensitivity settings for different features can be adjusted and customized; validation by an experienced radiologist co-reader is warranted [69].

AI tools may be useful in practice settings with limited subspeciality expertise: using AI solutions in the settings with minimal radiology support and high negative predictive value may provide comfort for clinicians with no need for follow-up of benign findings, although this should be addressed with caution. Depending on the institutional cohorts, AI results cannot be generalized, as it is assumed that AI would misperform in specialized centers with higher malignancy rates in comparison to the average population [69]. In terms of legal frame, AI-generated conclusions being reviewed by board-certified radiologists or US practitioners, regardless of their specialty, is mandatory. Several authors suggest use of AI results as second-opinion, although this has a negative impact on workflow speed [10, 11,69]. US practices, in conjunction with vendors, should implement AI performance and quality control protocols in order to assess the reliability of the tool.

Finally, a limitation of AI should be noted: thyroid US scanning includes comprehensive neck soft tissue assessment, including lymph nodes and parathyroid glands, but currently, AI solutions address only one aspect of this complex examination.

5. Conclusions

The introduction of AI was a revolutionary event in thyroid nodule assessment. Not only ultrasound, but also other imaging methods such as CT and MRI, use it effectively [70–72]. In some cases, there is even the possibility to effectively predict the immunohistochemistry of the thyroid nodule simply through the evaluation of segmented image datasets by AI systems [73]. Moreover, the use of CAD in daily clinical practice does not have a significant impact on workflow, as it increases the examination time by approximately 2–3 min [4]

However, the real effectiveness of AI systems remains controversial; taking into consideration the largest and most scientifically valid studies, it is possible to state that AI provides results that are comparable or in any case inferior to that of expert radiologists. Furthermore, it is necessary to consider the relevant heterogeneity of sensitivity and specificity between studies, due to the diversity in methodology and to the differences among patients included [39].

AI systems still have a long way to go to replace experienced radiologists in the process of improving accuracy and reducing time consumption, and larger studies meeting uniformity criteria are necessary to evaluate the diagnostic performance of these systems further. Nevertheless, the current CAD systems offer support for radiologists in thyroid nodule assessment and increase the overall accuracy in routine thyroid US [10,11,39].

AI solutions with CAD should be implemented in the teaching process of junior specialists. Deep-learning algorithms would benefit from follow-up US imaging data of the same thyroid nodules in combination with TIRADS classification, rather than dichotomous prediction, to increase their repeatability, reliability, and accuracy.

Regarding the legal frame, AI-generated conclusions should be reviewed by board-certified radiologists or US practitioners as mandatory practice, such that AI results may be provided only as a second opinion.

References

1. Curado, M.P.; Edwards, B.; Shin, H.R.; Storm, H.; Ferlay, J.; Heanue, M.; Boyle, P. *Cancer Incidence in Five Continents*; Iarc Scientific Publications: Lyon, France, 2014; Volume 10.
2. Grani, G.; Zatelli, M.C.; Alfò, M.; Montesano, T.; Torlontano, M.; Morelli, S.; Deandrea, M.; Antonelli, A.; Francese, C.; Ceresini, G.; et al. Real-World Performance of the American Thyroid Association Risk Estimates in Predicting 1-Year Differentiated Thyroid Cancer Outcomes: A Prospective Multicenter Study of 2000 Patients. *Thyroid* **2021**, *31*, 264–271. [CrossRef] [PubMed]
3. Celletti, I.; Fresilli, D.; De Vito, C.; Bononi, M.; Cardaccio, S.; Cozzolino, A.; Durante, C.; Grani, G.; Grimaldi, G.; Isidori, A.M.; et al. TIRADS, SRE and SWE in INDETERMINATE thyroid nodule characterization: Which has better diagnostic performance? *Radiol. Med.* **2021**, *126*, 1189–1200. [CrossRef] [PubMed]
4. Fresilli, D.; Grani, G.; De Pascali, M.L.; Alagna, G.; Tassone, E.; Ramundo, V.; Ascoli, V.; Bosco, D.; Biffoni, M.; Bononi, M.; et al. Computer-aided diagnostic system for thyroid nodule sonographic evaluation outperforms the specificity of less experienced examiners. *J. Ultrasound* **2020**, *23*, 169–174. [CrossRef] [PubMed]
5. Gabillard, J.C.; Ulisse, S.; Baldini, E.; Sorrenti, S.; Cremet, J.Y.; Coccaro, C.; Prigent, C.; D'Armiento, M.; Arlot-Bonnemains, Y. Aurora-C interacts with and phosphorylates the transforming acidic coiled-coil 1 protein. *Biochem. Biophys. Res. Commun.* **2011**, *408*, 647–653. [CrossRef] [PubMed]
6. Baldini, E.; Tuccilli, C.; Prinzi, N.; Sorrenti, S.; Falvo, L.; De Vito, C.; Catania, A.; Tartaglia, F.; Mocini, R.; Coccaro, C.; et al. Deregulated expression of Aurora kinases is not a prognostic biomarker in papillary thyroid cancer patients. *PLoS ONE* **2015**, *10*, e0121514. [CrossRef]
7. Beam, A.L.; Kohane, I.S. Big Data and Machine Learning in Health Care. *JAMA* **2018**, *319*, 1317–1318. [CrossRef]
8. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **2016**, *278*, 563–577. [CrossRef]
9. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006.
10. Cantisani, V.; Grani, G.; Tovoli, F.; Piscaglia, F.; Catalano, C. Artificial Intelligence: What Is It and How Can It Expand the Ultrasound Potential in the Future? *Ultraschall Med.* **2020**, *41*, 356–360. [CrossRef]
11. Thomas, J.; Ledger, G.A.; Mamillapalli, C.K. Use of artificial intelligence and machine learning for estimating malignancy risk of thyroid nodules. *Curr. Opin. Endocrinol. Diabetes Obes.* **2020**, *27*, 345–350. [CrossRef]
12. Bini, F.; Pica, A.; Azzimonti, L.; Giusti, A.; Ruinelli, L.; Marinozzi, F.; Trimboli, P. Artificial Intelligence in Thyroid Field—A Comprehensive Review. *Cancers* **2021**, *13*, 4740. [CrossRef]
13. Sorrenti, S.; Dolcetti, V.; Fresilli, D.; Del Gaudio, G.; Pacini, P.; Huang, P.; Camponovo, C.; Leoncini, A.; D'Andrea, V.; Pironi, D.; et al. The Role of CEUS in the Evaluation of Thyroid Cancer: From Diagnosis to Local Staging. *J. Clin. Med.* **2021**, *10*, 4559. [CrossRef] [PubMed]
14. Zhu, Y.C.; Jin, P.F.; Bao, J.; Jiang, Q.; Wang, X. Thyroid ultrasound image classification using a convolutional neural network. *Ann. Transl Med.* **2021**, *9*, 1526. [CrossRef] [PubMed]
15. Peng, S.; Liu, Y.; Lv, W.; Liu, L.; Zhou, Q.; Yang, H.; Ren, J.; Liu, G.; Wang, X.; Zhang, X.; et al. Deep learning-based artificial intelligence model to assist thyroid nodule diagnosis and management: A multicentre diagnostic study. *Lancet Digit. Health* **2021**, *3*, e250–e259. [CrossRef]
16. Bai, Z.; Chang, L.; Yu, R.; Li, X.; Wei, X.; Yu, M.; Liu, Z.; Gao, J.; Zhu, J.; Zhang, Y.; et al. Thyroid nodules risk stratification through deep learning based on ultrasound images. *Med. Phys.* **2020**, *47*, 6355–6365. [CrossRef]

17. Yoon, J.; Lee, E.; Kang, S.W.; Han, K.; Park, V.Y.; Kwak, J.Y. Implications of US radiomics signature for predicting malignancy in thyroid nodules with indeterminate cytology. *Eur. Radiol.* **2021**, *31*, 5059–5067. [CrossRef]

18. Liu, Z.; Zhong, S.; Liu, Q.; Xie, C.; Dai, Y.; Peng, C.; Chen, X.; Zou, R. Thyroid nodule recognition using a joint convolutional neural network with information fusion of ultrasound images and radiofrequency data. *Eur. Radiol.* **2021**, *31*, 5001–5011. [CrossRef] [PubMed]

19. Gomes Ataide, E.J.; 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*, 6110. [CrossRef]

20. Ye, H.; Hang, J.; Chen, X.; Di Xu Chen, J.; Ye, X.; Zhang, D. An intelligent platform for ultrasound diagnosis of thyroid nodules. *Sci. Rep.* **2020**, *10*, 13223. [CrossRef]

21. Wei, X.; Gao, M.; Yu, R.; Liu, Z.; Gu, Q.; Liu, X.; Zheng, Z.; Zheng, X.; Zhu, J.; Zhang, S. Ensemble Deep Learning Model for Multicenter Classification of Thyroid Nodules on Ultrasound Images. *Med. Sci. Monit.* **2020**, *26*, e926096. [CrossRef]

22. Sun, C.; Zhang, Y.; Chang, Q.; Liu, T.; Zhang, S.; Wang, X.; Guo, Q.; Yao, J.; Sun, W.; Niu, L. Evaluation of a deep learning-based computer-aided diagnosis system for distinguishing benign from malignant thyroid nodules in ultrasound images. *Med. Phys.* **2020**, *47*, 3952–3960. [CrossRef]

23. Zhou, H.; Jin, Y.; Dai, L.; Zhang, M.; Qiu, Y.; Wang, K.; Tian, J.; Zheng, J. Differential Diagnosis of Benign and Malignant Thyroid Nodules Using Deep Learning Radiomics of Thyroid Ultrasound Images. *Eur. J. Radiol.* **2020**, *127*, 108992. [CrossRef] [PubMed]

24. Nguyen, D.T.; Kang, J.K.; Pham, T.D.; Batchuluun, G.; Park, K.R. Ultrasound Image-Based Diagnosis of Malignant Thyroid Nodule Using Artificial Intelligence. *Sensors* **2020**, *20*, 1822. [CrossRef] [PubMed]

25. Wang, L.; Zhang, L.; Zhu, M.; Qi, X.; Yi, Z. Automatic diagnosis for thyroid nodules in ultrasound images by deep neural networks. *Med. Image Anal.* **2020**, *61*, 101665. [CrossRef] [PubMed]

26. Thomas, J.; Haertling, T. AIBx, Artificial Intelligence Model to Risk Stratify Thyroid Nodules. *Thyroid* **2020**, *30*, 878–884. [CrossRef] [PubMed]

27. Galimzianova, A.; Siebert, S.M.; Kamaya, A.; Rubin, D.L.; Desser, T.S. Quantitative Framework for Risk Stratification of Thyroid Nodules With Ultrasound: A Step Toward Automated Triage of Thyroid Cancer. *AJR Am. J. Roentgenol.* **2020**, *214*, 885–892. [CrossRef] [PubMed]

28. Nguyen, D.T.; Pham, T.D.; Batchuluun, G.; Yoon, H.S.; Park, K.R. Artificial Intelligence-Based Thyroid Nodule Classification Using Information from Spatial and Frequency Domains. *J. Clin. Med.* **2019**, *8*, 1976. [CrossRef] [PubMed]

29. Buda, M.; Wildman-Tobriner, B.; Hoang, J.K.; Thayer, D.; Tessler, F.N.; Middleton, W.D.; Mazurowski, M.A. Management of Thyroid Nodules Seen on US Images: Deep Learning May Match Performance of Radiologists. *Radiology* **2019**, *292*, 695–701. [CrossRef]

30. Koh, J.; Lee, E.; Han, K.; Kim, E.K.; Son, E.J.; Sohn, Y.M.; Seo, M.; Kwon, M.R.; Yoon, J.H.; Lee, J.H.; et al. Diagnosis of thyroid nodules on ultrasonography by a deep convolutional neural network. *Sci. Rep.* **2020**, *10*, 15245. [CrossRef]

31. Wang, L.; Yang, S.; Yang, S.; Zhao, C.; Tian, G.; Gao, Y.; Chen, Y.; Lu, Y. Automatic thyroid nodule recognition and diagnosis in ultrasound imaging with the YOLOv2 neural network. *World J. Surg. Oncol.* **2019**, *17*, 12. [CrossRef]

32. Han, M.; Ha, E.J.; Park, J.H. Computer-Aided Diagnostic System for Thyroid Nodules on Ultrasonography: Diagnostic Performance Based on the Thyroid Imaging Reporting and Data System Classification and Dichotomous Outcomes. *AJNR Am. J. Neuroradiol.* **2021**, *42*, 559–565. [CrossRef]

33. Zhang, Y.; Wu, Q.; Chen, Y.; Wang, Y. A Clinical Assessment of an Ultrasound Computer-Aided Diagnosis System in Differentiating Thyroid Nodules With Radiologists of Different Diagnostic Experience. *Front. Oncol.* **2020**, *10*, 557169. [CrossRef] [PubMed]

34. Jin, Z.; Zhu, Y.; Zhang, S.; Xie, F.; Zhang, M.; Zhang, Y.; Tian, X.; Zhang, J.; Luo, Y.; Cao, J. Ultrasound Computer-Aided Diagnosis (CAD) Based on the Thyroid Imaging Reporting and Data System (TI-RADS) to Distinguish Benign from Malignant Thyroid Nodules and the Diagnostic Performance of Radiologists with Different Diagnostic Experience. *Med. Sci. Monit.* **2020**, *26*, e918452. [CrossRef] [PubMed]

35. Xia, S.; Yao, J.; Zhou, W.; Dong, Y.; Xu, S.; Zhou, J.; Zhan, W. A computer-aided diagnosing system in the evaluation of thyroid nodules-experience in a specialized thyroid center. *World J. Surg. Oncol.* **2019**, *17*, 210. [CrossRef] [PubMed]

36. Jin, A.; Li, Y.; Shen, J.; Zhang, Y.; Wang, Y. Clinical Value of a Computer-Aided Diagnosis System in Thyroid Nodules: Analysis of a Reading Map Competition. *Ultrasound Med. Biol.* **2019**, *45*, 2666–2671. [CrossRef]

37. Kim, H.L.; Ha, E.J.; Han, M. Real-World Performance of Computer-Aided Diagnosis System for Thyroid Nodules Using Ultrasonography. *Ultrasound Med. Biol.* **2019**, *45*, 2672–2678. [CrossRef]

38. Chi, J.; Walia, E.; Babyn, P.; Wang, J.; Groot, G.; Eramian, M. Thyroid Nodule Classification in Ultrasound Images by Fine-Tuning Deep Convolutional Neural Network. *J. Digit. Imaging* **2017**, *30*, 477–486. [CrossRef]

39. Zhao, W.J.; Fu, L.R.; Huang, Z.M.; Zhu, J.Q.; Ma, B.Y. Effectiveness evaluation of computer-aided diagnosis system for the diagnosis of thyroid nodules on ultrasound: A systematic review and meta-analysis. *Medicine* **2019**, *98*, e16379. [CrossRef]

40. Watkins, L.; O'Neill, G.; Young, D.; McArthur, C. 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. *Br. J. Radiol.* **2021**, *94*, 20201444. [CrossRef]

41. Wang, S.; Xu, J.; Tahmasebi, A.; Daniels, K.; Liu, J.B.; Curry, J.; Cottrill, E.; Lyshchik, A.; Eisenbrey, J.R. Incorporation of a Machine Learning Algorithm With Object Detection Within the Thyroid Imaging Reporting and Data System Improves the Diagnosis of Genetic Risk. *Front. Oncol.* **2020**, *10*, 591846. [CrossRef]

42. Wildman-Tobriner, B.; Buda, M.; Hoang, J.K.; Middleton, W.D.; Thayer, D.; Short, R.G.; Tessler, F.N.; Mazurowski, M.A. Using Artificial Intelligence to Revise ACR TI-RADS Risk Stratification of Thyroid Nodules: Diagnostic Accuracy and Utility. *Radiology* **2019**, *292*, 112–119. [CrossRef]

43. Lambin, P.; Rios-Velazquez, E.; Leijenaar, R.; Carvalho, S.; van Stiphout, R.G.; Granton, P.; Zegers, C.M.; Gillies, R.; Boellard, R.; Dekker, A.; et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **2012**, *48*, 441–446. [CrossRef] [PubMed]

44. Sollini, M.; Cozzi, L.; Chiti, A.; Kirienko, M. Texture analysis and machine learning to characterize suspected thyroid nodules and differentiated thyroid cancer: Where do we stand? *Eur. J. Radiol.* **2018**, *99*, 1–8. [CrossRef]

45. Han, Z.; Feng, N.; Lu, Y.; Li, M.; Wei, P.; Yao, J.; Zhu, Q.; Lei, Z.; Xu, D. A Control Study on the Value of the Ultrasound Grayscale Ratio for the Differential Diagnosis of Thyroid Micropapillary Carcinoma and Micronodular Goiter in Two Medical Centers. *Front. Oncol.* **2021**, *10*, 625238. [CrossRef]

46. Grani, G.; D'Alessandri, M.; Carbotta, G.; Nesca, A.; Del Sordo, M.; Alessandrini, S.; Coccaro, C.; Rendina, R.; Bianchini, M.; Prinzi, N.; et al. Grey-Scale Analysis Improves the Ultrasonographic Evaluation of Thyroid Nodules. *Medicine* **2015**, *94*, e1129. [CrossRef]

47. Lei, Z.K.; Li, M.K.; Luo, D.C.; Han, Z.J. The clinical significance of ultrasound grayscale ratio in differentiating markedly hypoechoic and anechoic minimal thyroid nodules. *J. Cancer Res. Ther.* **2018**, *14*, 1567–1571. [PubMed]

48. Chen, X.; Gao, M.; Hu, L.; Zhu, J.; Zhang, S.; Wei, X. The diagnostic value of the ultrasound gray scale ratio for different sizes of thyroid nodules. *Cancer Med.* **2019**, *8*, 7644–7649. [CrossRef] [PubMed]

49. Han, Z.; Lei, Z.; Li, M.; Luo, D.; Ding, J. Differential diagnosis value of the ultrasound gray scale ratio for papillary thyroid microcarcinomas and micronodular goiters. *Quant. Imaging Med. Surg.* **2018**, *8*, 507–513. [CrossRef]

50. Liang, J.; Huang, X.; Hu, H.; Liu, Y.; Zhou, Q.; Cao, Q.; Wang, W.; Liu, B.; Zheng, Y.; Li, X.; et al. Predicting Malignancy in Thyroid Nodules: Radiomics Score Versus 2017 American College of Radiology Thyroid Imaging, Reporting and Data System. *Thyroid* **2018**, *28*, 1024–1033. [CrossRef] [PubMed]

51. Kwon, M.R.; Shin, J.H.; Park, H.; Cho, H.; Hahn, S.Y.; Park, K.W. Radiomics Study of Thyroid Ultrasound for Predicting BRAF Mutation in Papillary Thyroid Carcinoma: Preliminary Results. *AJNR Am. J. Neuroradiol.* **2020**, *41*, 700–705. [CrossRef]

52. Park, V.Y.; Han, K.; Kim, H.J.; Lee, E.; Youk, J.H.; Kim, E.K.; Moon, H.J.; Yoon, J.H.; Kwak, J.Y. Radiomics signature for prediction of lateral lymph node metastasis in conventional papillary thyroid carcinoma. *PLoS ONE* **2020**, *15*, e0227315. [CrossRef]

53. Akkus, Z.; Cai, J.; Boonrod, A.; Zeinoddini, A.; Weston, A.D.; Philbrick, K.A.; Erickson, B.J. A Survey of Deep-Learning Applications in Ultrasound: Artificial Intelligence-Powered Ultrasound for Improving Clinical Workflow. *J. Am. Coll. Radiol.* **2019**, *16*, 1318–1328. [CrossRef] [PubMed]

54. Yoon, J.; Lee, E.; Koo, J.S.; Yoon, J.H.; Nam, K.H.; Lee, J.; Jo, Y.S.; Moon, H.J.; Park, V.Y.; Kwak, J.Y. Artificial intelligence to predict the BRAFV600E mutation in patients with thyroid cancer. *PLoS ONE* **2020**, *15*, e0242806. [CrossRef] [PubMed]

55. Ulisse, S.; Baldini, E.; Lauro, A.; Pironi, D.; Tripodi, D.; Lori, E.; Ferent, I.C.; Amabile, M.I.; Catania, A.; Di Matteo, F.M.; et al. Papillary Thyroid Cancer Prognosis: An Evolving Field. *Cancers* **2021**, *13*, 5567. [CrossRef] [PubMed]

56. Baldini, E.; Tuccilli, C.; Pironi, D.; Catania, A.; Tartaglia, F.; Di Matteo, F.M.; Palumbo, P.; Arcieri, S.; Mascagni, D.; Palazzini, G.; et al. Expression and Clinical Utility of Transcription Factors Involved in Epithelial-Mesenchymal Transition during Thyroid Cancer Progression. *J. Clin. Med.* **2021**, *10*, 4076. [CrossRef] [PubMed]

57. Daniels, K.; Gummadi, S.; Zhu, Z.; Wang, S.; Patel, J.; Swendseid, B.; Lyshchik, A.; Curry, J.; Cottrill, E.; Eisenbrey, J. Machine Learning by Ultrasonography for Genetic Risk Stratification of Thyroid Nodules. *JAMA Otolaryngol. Head Neck Surg.* **2020**, *146*, 36–41. [CrossRef]

58. Prieditis, P.; Radzina, M.; Mikijanska, M.; Liepa, M.; Stepanovs, K.; Grani, G.; Durante, C.; Lamartina, L.; Trimboli, P.; Cantisani, V. Non-Marked Hypoechogenic Nodules: Multicenter Study on the Thyroid Malignancy Risk Stratification and Accuracy Based on TIRADS Systems Comparison. *Medicina* **2022**, *58*, 257. [CrossRef]

59. Gild, M.L.; Chan, M.; Gajera, J.; Lurie, B.; Gandomkar, Z.; Clifton-Bligh, R.J. Risk stratification of indeterminate thyroid nodules using ultrasound and machine learning algorithms. *Clin. Endocrinol.* **2022**, *96*, 646–652. [CrossRef] [PubMed]

60. Chen, Y.; Gao, Z.; He, Y.; Mai, W.; Li, J.; Zhou, M.; Li, S.; Yi, W.; Wu, S.; Bai, T.; et al. An Artificial Intelligence Model Based on ACR TI-RADS Characteristics for US Diagnosis of Thyroid Nodules. *Radiology* **2022**, *22*, 211455. [CrossRef]

61. Xu, L.; Gao, J.; Wang, Q.; Yin, J.; Yu, P.; Bai, B.; Pei, R.; Chen, D.; Yang, G.; Wang, S.; et al. Computer-Aided Diagnosis Systems in Diagnosing Malignant Thyroid Nodules on Ultrasonography: A Systematic Review and Meta-Analysis. *Eur. Thyroid J.* **2020**, *9*, 186–193. [CrossRef]

62. Elliott Range, D.D.; Dov, D.; Kovalsky, S.Z.; Henao, R.; Carin, L.; Cohen, J. Application of a machine learning algorithm to predict malignancy in thyroid cytopathology. *Cancer Cytopathol.* **2020**, *128*, 287–295. [CrossRef]

63. Fresilli, D.; David, E.; Pacini, P.; Del Gaudio, G.; Dolcetti, V.; Lucarelli, G.T.; Di Leo, N.; Bellini, M.I.; D'Andrea, V.; Sorrenti, S.; et al. Thyroid Nodule Characterization: How to Assess the Malignancy Risk. Update of the Literature. *Diagnostics* **2021**, *11*, 1374. [CrossRef]

64. Friedrich-Rust, M.; Meyer, G.; Dauth, N.; Berner, C.; Bogdanou, D.; Herrmann, E.; Zeuzem, S.; Bojunga, J. Interobserver agreement of Thyroid Imaging Reporting and Data System (TIRADS) and strain elastography for the assessment of thyroid nodules. *PLoS ONE* **2013**, *8*, e77927. [CrossRef] [PubMed]

65. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]

66. Kezlarian, B.; Lin, O. Artificial Intelligence in Thyroid Fine Needle Aspiration Biopsies. *Acta Cytol.* **2021**, *65*, 324–329. [CrossRef]

67. Bellini, M.I.; Biffoni, M.; Patrone, R.; Borcea, M.C.; Costanzo, M.L.; Garritano, T.; Melcarne, R.; Menditto, R.; Metere, A.; Scorziello, C.; et al. Poorly Differentiated Thyroid Carcinoma: Single Centre Experience and Review of the Literature. *J. Clin. Med.* **2021**, *10*, 5258. [CrossRef]

68. Ippolito, A.M.; De Laurentiis, M.; La Rosa, G.L.; Eleuteri, A.; Tagliaferri, R.; De Placido, S.; Vigneri, R.; Belfiore, A. Neural network analysis for evaluating cancer risk in thyroid nodules with an indeterminate diagnosis at aspiration cytology: Identification of a low-risk subgroup. *Thyroid* **2004**, *14*, 1065–1071. [CrossRef]

69. Wildman-Tobriner, B.; Taghi-Zadeh, E.; Mazurowski, M.A. Artificial Intelligence (AI) Tools for Thyroid Nodules on Ultrasound, From the AJR Special Series on AI Applications. *AJR Am. J. Roentgenol.* **2022**, 1–8. [CrossRef] [PubMed]

70. Li, W.; Cheng, S.; Qian, K.; Yue, K.; Liu, H. Automatic Recognition and Classification System of Thyroid Nodules in CT Images Based on CNN. *Comput. Intell. Neurosci.* **2021**, *2021*, 5540186. [CrossRef]

71. Zhao, H.B.; Liu, C.; Ye, J.; Chang, L.F.; Xu, Q.; Shi, B.W.; Liu, L.L.; Yin, Y.L.; Shi, B.B. A comparison between deep learning convolutional neural networks and radiologists in the differentiation of benign and malignant thyroid nodules on CT images. *Endokrynol. Pol.* **2021**, *72*, 217–225. [CrossRef] [PubMed]

72. Wang, H.; Song, B.; Ye, N.; Ren, J.; Sun, X.; Dai, Z.; Zhang, Y.; Chen, B.T. Machine learning-based multiparametric MRI radiomics for predicting the aggressiveness of papillary thyroid carcinoma. *Eur. J. Radiol.* **2020**, *122*, 108755. [CrossRef]

73. Gu, J.; Zhu, J.; Qiu, Q.; Wang, Y.; Bai, T.; Yin, Y. Prediction of Immunohistochemistry of Suspected Thyroid Nodules by Use of Machine Learning-Based Radiomics. *AJR Am. J. Roentgenol.* **2019**, *213*, 1348–1357. [CrossRef] [PubMed]