A systematic review on the role of artificial intelligence in sonographic diagnosis of thyroid cancer: Past, present and future

Fatemeh Abdolali¹, ², Atefeh Shahroudnejad¹, ², Abhilash Rakkunedeth Hareendranathan², Jacob L Jaremko¹, ², Michelle Noga³ Kumaradevan Punithakumar³

¹ Department of Radiology and Diagnostic Imaging, University of Alberta Hospital, Edmonton, Alberta, Canada, T6G2B7
² MEDO.ai, Dual Headquarters at Singapore and Edmonton, Alberta, Canada, T5J4P6
³ E-mail: abdolali@ualberta.ca

Abstract: Thyroid cancer is common worldwide, with a rapid increase in prevalence across North America in recent years. While most patients present with palpable nodules through physical examination, a large number of small and medium-sized nodules are detected by ultrasound examination. Suspicious nodules are then sent for biopsy through fine needle aspiration. Since biopsies are invasive and sometimes inconclusive, various research groups have tried to develop computer-aided diagnosis systems. Earlier approaches along these lines relied on clinically relevant features that were manually identified by radiologists. With the recent success of artificial intelligence (AI), various new methods are being developed to identify these features in thyroid ultrasound automatically. In this paper, we present a systematic review of state-of-the-art on AI application in sonographic diagnosis of thyroid cancer. This review follows a methodology-based classification of the different techniques available for thyroid cancer diagnosis. With more than 50 papers included in this review, we reflect on the trends and challenges of the field of sonographic diagnosis of thyroid malignancies and potential of computer-aided diagnosis to increase the impact of ultrasound applications on the future of thyroid cancer diagnosis. Machine learning will continue to play a fundamental role in the development of future thyroid cancer diagnosis frameworks.

1 Introduction

The incidence of thyroid cancer has increased at an alarming rate in the United States with 52,070 cases diagnosed in 2019, out of which 2170 cases resulted in death [1]. It is the most common cancer in American women aged between 20 and 34 [1]. The corresponding numbers in Canada are equally significant with the Canadian Cancer Statistics estimating that 8,200 Canadians (6100 women and 2100 men) would be diagnosed with thyroid cancer resulting in 230 deaths in 2019 [2]. The incidence of thyroid cancer has steadily increased from 1970, with the most substantial increase in middle-aged women [3].

Ultrasound (US) is a fast, safe and inexpensive imaging technique that can provide a complete visualization of thyroid nodules. Imaging is generally performed in both transverse and sagittal orientation using ultrasound probes in the 7 - 15 MHz range. Thyroid nodules have several distinguishable characteristics on ultrasound such as shape, size, echogenicity (brightness) and echotexture (composition). For example, features such as microcalcification or taller-than-wide nodule shape can be predictors of malignancy, whereas spongiform appearance can be representative of benign case [4].

The critical challenge in image aided thyroid nodule diagnosis and reporting of thyroid nodules is the extraction of optimal sets of features from ultrasound images that differentiate malignant from benign nodules. In order to standardize the reporting of thyroid nodules, American College of Radiology introduced a Thyroid Imaging, Reporting and Data System (TIRADS) [5]. The goal of TIRADS is to provide standardized recommendations for the management of thyroid nodules on the basis of well-defined clinical features for every lesion. TIRADS identifies six categories for a given nodule: composition, echogenicity, shape, size, margins and echogenic foci. Based on the scores of each of these features, the nodule is classified as benign, minimally suspicious, moderately suspicious or highly suspicious of malignancy. Various studies have reported increased...
agreement among users while using TIRADS. A parallel approach aimed at reducing the variability in thyroid nodule reporting is based on computer-aided diagnosis (CAD) systems that usually categorize nodules as malignant or benign. These systems are mostly trained on retrospectively collected biopsy data and generally aim to reduce false positives that would otherwise result in unnecessary biopsies and over-diagnosis [6].

In this paper, we review various artificial intelligence (AI)-based CAD systems that have been proposed for thyroid ultrasound nodule analysis, including segmentation, detection and classification. Earlier reviews have focused on specific methods such as feature-based [7], deep learning [8], linear [9] and non-linear [10] approaches (see Table 1). We aim to provide a broader perspective in our discussion and compare the above-mentioned approaches to one another. Our review also includes more recent publications in each of these approaches. This paper presents, for the first time, a systematic review of the existing literature on sonographic diagnosis of thyroid nodules, which covers both classical and deep learning methods.

Figure 1 shows the histogram of the papers introducing novel machine learning-based techniques for thyroid nodule diagnosis. Inclusion and exclusion criteria are summarized in Table 2. With more than 50 articles included in this survey, we focus on the methodological aspects of those works. We propose a structured analysis of the different approaches used for ultrasound-based thyroid tissue quantification. The paper is organized as follows. Section 2 presents an overview of the parameterization models for thyroid ultrasound nodule analysis. In Section 3, we will discuss the main findings of previous works, limitations, challenges and future trends that are specific to thyroid US images. Conclusions are presented in Section 4.

![Figure 1: Publications on sonographic diagnosis of thyroid per year based on PubMed and Scopus search engines.](image)

![Table 1: Overview of review articles in diagnosis of thyroid cancer](table)

| Article | Year | Study purpose | Imaging modalities | Number of papers |
|---------|------|---------------|--------------------|------------------|
| Acharya *et al.* [7] | 2014 | Summarizing studies on malignancy detection | US | 57 |
| Khachnaoui *et al.* [8] | 2018 | Summarizing deep learning-based studies on thyroid diagnosis | US | 8 |
| Sollini *et al.* [9] | 2018 | Summarizing studies on texture analysis with and without CAD | US, CT, MRI, PET | 66 |
| Verburg *et al.* [10] | 2019 | Meta-analysis study for US classification of thyroid cancer | US | 10 |

## 2 Review of methodological approaches for thyroid ultrasound analysis

The aim of this section is to provide a structured reference guide for the different techniques of thyroid ultrasound analysis from a methodological point of view. All these works share a common goal: use CAD for thyroid parameterization. Figure 2 shows the general block diagram of a CAD system for sonographic diagnosis of the thyroid cancer. At first, some preprocessing steps such as rescaling pixel intensities and noise reduction are performed and region-of-interest (ROI) is selected from every US image. Then, feature extraction approach is applied. Finally, thyroid gland or nodule region is segmented so that the prediction about the malignancy could be performed. To find relevant literature sources, PubMed
and Scopus search engines were utilized using the query string: (a) “AI artificial intelligence” or “AI deep learning” or “AI segmentation” or “AI detection” and (b) “AI thyroid.” The last search was conducted on 12 December 2019. The search produced 997 results from search queries. Titles and abstracts were reviewed to determine whether they were suitable for this review. We adopted a set of inclusion and exclusion criteria (see Table 2). The articles were downloaded, and the titles, abstract and methods sections were reviewed in order to verify that the articles met the aforementioned criteria. A total of 52 articles met the criteria for inclusion in the final review. Overall, our review was aimed at describing recent artificial intelligence achievements in sonographic diagnosis of thyroid nodules.

Based on the papers surveyed, we identified 2 major categories: 1) clinical features-based models; 2) machine learning-based models; and several other sub-categories of methods.

### Table 2: Inclusion and exclusion criteria

| Inclusion criteria | Exclusion criteria |
|--------------------|--------------------|
| (a) Published between 2009 and 2019. | (a) Studies that utilized imaging modalities other than US. |
| (b) Peer-reviewed article. | (b) Studies that focused only on case reports. |

**Fig. 2**: Block diagram of a typical CAD system for sonographic diagnosis of thyroid cancer

**Fig. 3**: Transverse US image showing a thyroid nodule in the right thyroid lobe and adjacent anatomical structures.

### 2.1 Clinical features-based Models

Figure 3 shows an example of an image of the thyroid gland obtained by ultrasound. As can be appreciated, the correct interpretation of ultrasound images requires an expert radiologist. Zahang et al. [11] performed malignancy detection based on both conventional US and real-time elastography (RTE). A set of clinical features, i.e., echogenicity, margins, internal composition, aspect ratio, vascularity, hypoechoic halo, calcifications and real-time elastography grade were fed to nine different conventional classifiers. It was concluded that random field classifier based on US and RTE features outperformed other classifiers. In [12], ReliefF feature selection was utilized to select the most discriminative sonographic features. Selected features were fed into extreme learning machine (ELM), support vector machine (SVM) and neural network. In another research [13], sonographic features were fed into statistical classifiers to differentiate malignant thyroid nodules.

### 2.2 Machine learning-based models

Several promising machine learning-based models were proposed for sonographic diagnosis of thyroid cancer. Based on the papers surveyed, we identified 2 major categories: 1) classical features-based models; 2) Deep learning models, as discussed below.

#### 2.2.1 Classical features-based models

The use of classical features has been extensively studied in a variety of applications, including automatic segmentation of thyroid and nodules, thyroid cancer detection and classification. Their ability to integrate knowledge about location, size and shape of the thyroid organ or nodule has led to some of the promising segmentation [14] and detection methods [12, 19, 20]. Early classical features-based frameworks used discriminative statistical and texture features, i.e., wavelet transformation [21, 22], histogram feature [18], block difference of inverse probabilities (BDIP) and normalized multiscale intensity difference (NMISD) features [18] and co-occurrence matrix [21]. Then, feature selection techniques, i.e., minimum redundancy maximum
relevance (mRMR) [23] and T-test [24] were utilized to select those features which contribute most to the model. The selected feature set was fed to classifier, i.e., SVM [12, 19, 20, 23], Random forest [11, 15, 19], neural network [11, 12, 22], or Adaboost [11, 21]. These approaches can be divided into thyroid or nodule segmentation and thyroid cancer detection, as discussed below.

**Thyroid and nodule segmentation:** Several methods for automated segmentation of thyroid or nodules were proposed in the literature. Illanes et al. utilized a combination of wavelet texture extraction and parametrical modelling for US thyroid segmentation [14]. In [15], an algorithm based on iterative random walks solver and gradient vector flow based interframe belief propagation technique was developed to segment thyroid, lumen and external elastic laminae. Koundal et al. [16] proposed Spatial Neutrosophic Distance Regularized Level Set method based on Neutrosophic L-Means clustering to segment thyroid nodules. Later, they used Neutrosophic Nakagami Total Variation and small rectangular region extraction as pre-processing steps [27] to improve the segmentation results. Multi-organ segmentation of thyroid, carotid artery, muscles and trachea was performed by Narayan et al. [17] using speckle patch similarity estimation. In [18], Radial Basis Function (RBF) neural network was trained to classify each block of US image into thyroid gland and non-thyroid gland pixels.

**Thyroid cancer detection and classification:** The use of classical features has been extensively studied for thyroid cancer detection. Two threshold binary decomposition was utilized in [19] for patch based feature extraction. The extracted features were fed into random forest and SVM classifiers. Yu et al. [20] developed a framework based on region-based active contours and texture features extraction. The best performance was achieved by combining Artificial neural networks (ANN) and SVM. In other work, Gabor features were extracted from high resolution US images [24]. Locality sensitive discriminant analysis was utilized for feature space reduction. The selected features were fed into SVM, k-nearest neighbors (KNN) and multilayer perceptron (MLP) classifiers. Wavelet transformation was utilized as feature extraction method in several studies [21, 22]. In another study, different linear and nonlinear machine learning algorithms were evaluated for classification of thyroid nodules [28].

**2.2.2 Deep learning models:** Deep neural networks are representation learning methods which extract features from input data and use these features to perform machine learning tasks such as segmentation or detection. These approaches can be divided into thyroid or nodule segmentation and thyroid cancer detection, as discussed below.

**Thyroid and nodule segmentation:** Poudel et al. [29] compared the performance of four popular segmentation methods, i.e., active contour without edges, graph cut, pixel based classifier and random forest classifier with 3D U-net [30] on OpenCAS thyroid dataset [31]. They concluded that 3D U-net outperforms other methods in terms of accuracy. Several deep learning architectures were utilized for nodule segmentation in the literature. In [32], mark-guided U-net [33] was proposed for segmentation of thyroid nodules. Four landmark points corresponding to the major and minor axes of a nodule were determined manually to guide the training and inference of U-Net. A deep CNN framework with multiple intermediate layers was introduced in [34] for thyroid nodule segmentation. The proposed CNN contains 15 convolutional layers and two pooling layers. The main drawback of the proposed approach is that some thyroid nodules with complicated back-ground cannot be accurately segmented. An 8 layer fully convolutional network was proposed in [35] for thyroid nodule segmentation. The proposed framework outperformed U-net on the local test dataset.

**Thyroid cancer detection and classification:** Several deep learning based frameworks were proposed by different groups of scientists in the recent years. A research group evaluated the accuracy of thyroid cancer diagnosis using deep learning [10]. Authors utilized the dataset from three different Chinese hospitals and concluded that deep learning approach resulted in slightly inferior sensitivity and substantially improved specificity. A deep learning approach inspired by faster R-CNN [36] was proposed in [37] for thyroid papillary cancer detection. Using layer concatenation strategy, more detailed features of low resolution US images were extracted. An end-to-end detection network based on integrating YOLOv2 [38] and Resnet v2-50 [39] was employed in [40]. A retrospective, multicohort, diagnostic study using ultrasound images sets from three hospitals in China was performed in [41]. The CNN model was developed by ensembling Resnet50 [42] and Darknet19 [43] models. Song et al. [44] proposed a deep learning framework based on Inception-V3 network model [45] for diagnosis of thyroid nodules. The model was trained on a small local dataset which consists of cropped nodules by a clinician. Two pre-trained CNN models, i.e., âÅÌimagenet-vgg-verydeep-16âÅÎ and âÅÌimagenet-vgg-fâÅÎ, were used in thyroid nodule malignancy detection [46]. To train the CNN with their local data, the ROIs were extracted by a radiologist from each US image.

**3 Discussion**

Our review of thyroid diagnosis techniques from US images has shown that accurate interpretation of ultrasound images is based on clinical feature-based models. However, recent advances in artificial
intelligence have introduced efficient and flexible machine learning techniques applied to the analysis of thyroid US images. Machine learning will continue to play a fundamental role in the development of future thyroid cancer diagnosis frameworks. In this section, we discuss the applications of artificial intelligence in sonographic diagnosis of thyroid cancer, as well as the limitations and opportunities in this growing field.

3.1 Analysis by reported results and main findings

Several CAD systems for sonographic diagnosis of thyroid cancer have been developed, from clinical feature-based systems to machine learning-based systems. See Table 5 and Table 3 for a detailed list of the publications reviewed in this paper. In this section, we compare the previous works based on the reported results. For nodule segmentation, several methods [16, 32, 35, 47] achieved Dice scores higher than 90%. The best performance was achieved using Spatial Neutrosophic Distance Regularized level set method based on Neutrosophic L-Means clustering [16]. For thyroid segmentation, several methods [14, 17, 18, 29] achieved Dices score higher than 85%. Chang et al. [18] proposed a framework based on Haar Wavelet features and RBF neural network to classify each block of images into thyroid gland and non-thyroid gland. They achieved mean Dice score of 96.52% for thyroid segmentation. Several groups utilized deep learning or classical features for thyroid nodule classification. In a recent study, spatial domain features based on deep learning and frequency domain features based on Fast Fourier transform were combined for classifying the input thyroid images into either benign or malignant cases [48]. The best classification accuracy was reported using multi-task cascade VGG-16 model [49].

The main limitations of previous works are discussed below.

3.2 Limitations, Challenges and Future Trends

During the last 2 decades, sonographic imaging of the thyroid gland was the most valued imaging method. However, there are several cases where the radiologist cannot distinguish between benign and malignant nodules with complete certainty. The reliability of diagnosis depends on the quality of US images and the expertise of the medical experts who interpret US images. Machine learning methods can help to extract features beyond human perception which leads to more effective characterization of thyroid anatomy.

Previous works focused on techniques for analysis of one nodule per US image. If one or more nodules are detected within the thyroid gland, it is not clear whether these techniques can be accurate enough. More experiments should be performed to evaluate the accuracy of previous techniques in such scenarios.

The limited availability of annotated ultrasound data has been a problem in automated sonographic diagnosis of thyroid cancer. This limitation becomes particularly relevant when implementing CAD systems for thyroid nodule classification, where large datasets are needed to characterize location and texture of thyroid gland. The need for large datasets is essential for the development and validation of new CAD systems for thyroid. Moreover, this represents a major obstacle to realize the full potential of deep learning based techniques. Although publicly available datasets with manual annotations of thyroid exist, the number of cases is limited to a couple of hundred cases at best. Public data repositories, i.e., DDTI [31] and Openicas dataset [50], provide open access to detailed manually-guided expert annotations of thyroid structure and nodules. Collection of a large comprehensive dataset is critical to develop future CAD systems that are robust to pathology.

4 Conclusions

In this paper, we presented the first systematic review of CAD systems for sonographic diagnosis of thyroid cancer. The continuing progress of machine learning have favored the development of complex and comprehensive models for sonographic diagnosis of thyroid nodules. As shown in this survey, the automatic parameterization of thyroid has been approached from different perspectives, (e.g. detection, segmentation and classification) using various methodologies. The categorization of approaches in this paper provides a reference guide to the current techniques available for the analysis of thyroid US images. We have also indicated current challenges and future opportunities in CAD systems for analysis of thyroid US images. New efficient machine learning models should embed the anatomical context inherent to the thyroid gland to provide the essential accuracy for clinical practice.

Acknowledgment

This work was supported by Mitacs through the Mitacs Accelerate Program (IT12871).

5 References

1 Siegel, R.L., Miller, K.D. and Jemal, A.: ‘Cancer statistics, 2019’, CA: a cancer journal for clinicians, 2019, 69, pp. 7–34
2 ‘Canadian Cancer Statistics Advisory Committee 2019’, http://www.cancer.ca/∼/media/cancer/ca/CW/cancer%202019/Canadian%20Cancer-Statistics-2018-EN.pdf?la=en
3 Grant, E.G., Tessler, F.N., Hoang, J.K., Langer, J.E., Beland, M.D., Berland, L.L., Cronan, J.J., Deser, T.S., Frates, M.C., Hamper, U.M. and Middleton, W.D.: ‘Thyroid ultrasound reporting lexicon: white paper of the ACR thyroid imaging, reporting and data system (TI-RADS) committee’, Journal of the American college of radiology, 2015, 12, pp. 1520–1533
4 Li, J., Ma, X. and Cui, K.: ‘ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee’, Journal of the American college of radiology, 2018, 15, pp. 380–381
an extreme learning machine approach', Computer methods and programs in biomedicine, 2017, 147, pp. 37–49
13 Patrício, M., Oliveira, C. and Caseiro-Alves, F.: ‘Differentiating malignant thyroid nodule with statistical classifiers based on demographic and ultrasound features’, Proc. IEEE 5th Portuguese Meeting on Bioengineering (ENBENG), Coimbra, Portugal, February 2017, pp. 1–4
14 Illanes, A., Esmaili, N., Poudel, P., Balakrishnan, S. and Friebe, M.: ‘Parametrical modelling for texture characterization\u2013A\u2013A novel approach applied to ultrasound thyroid segmentation’, PloS one, 2019, 14, e0211215
15 China, D., Illanes, A., Friebe, M., Mitra, P. and Sheet, D.: ‘Anatomical structure segmentation in ultrasound volumes using cross frame belief propagating iterative random walks’, IEEE journal of biomedical and health informatics, 2018, 23, pp. 1110–1118
16 Koundal, D., Gupta, S. and Singh, S.: ‘Automated delineation of thyroid nodules in ultrasound images using spatial neural network object and level set’, Applied Soft Computing, 2016, 40, pp. 86–97
17 Naryan, N.S., Marziliano, P., Kanagalangam, J. and Hobb, C.G.: ‘Speckle patch similarity for echogenicity-based multiorgan segmentation in ultrasound images of the thyroid gland’, IEEE journal of biomedical and health informatics, 2015, 21, pp. 172–183
18 Chang, C.Y., Lei, Y.F., Tseng, C.H. and Shih, S.R.: ‘Thyroid segmentation and volume estimation in ultrasound images’, Applied Soft Computing, 2010, 57, pp. 1348–1357

Table 3 Overview of applications of deep learning based methods in sonographic diagnosis of thyroid cancer

| Citation | Study Purpose | General information | Dataset | Method | Measures | Results or main findings | TEA/DS |
|----------|--------------|---------------------|---------|--------|----------|-------------------------|--------|
| 40-2017 | Thyroid nodule classification | Local training (259 benign cases, 257 malignant cases, test (55 with nodules- 23 normal cases) 276 fixated | 2D | Integration of ResNet-50 and U-Net | ROC | ROC=0.985 | ✓ |
| 41-2018 | Thyroid nodule classification | training (5125) validation (2000) | 2D | Resizable U-Net (ResNetV2 and DeepLabV5) | Sensitivity Specificity | ACU= 92.16% | ✓ |
| 42-2018 | Thyroid nodule classification | Not mentioned | 2D | Inception-V3 model | Sensitivity | Nodule positive predictive value | ✓ |
| 43-2018 | Thyroid nodule classification | Local training (1580) 478 benign, 444 malignant + 59 nodal Ki67 and + 40 thyroid biopsy test set | Not mentioned | Improved Faster R-CNN | TP/R, TP/R | AUC=95.5% | ✓ |
| 44-2018 | Thyroid nodule classification | Training data: 200 Test data: 80 | 2D | 8-layer self-supervised net (replace VGG features with in-house VGG-VGG without deconvolution layers) | Loss | ACU=98.66% | ✓ |
| 45-2018 | Thyroid nodule classification | 1148 images from Open-CAS dataset | 18 | Active Contour without Edges, Graph Cut, Pixel-Based Classifier, Random Forest Classifier and Convolutional Neural Network (CNN) | DiceAw coefficient | Hausdorff distance | ✓ |
| 46-2018 | Thyroid nodule classification | Local (653) | 2D | VGG-16 for extracting feature maps: classification using spatial pyramid method | AUC=95% | ✓ |
| 47-2018 | Thyroid nodule classification | Local (659) | 2D | U-Net: grid-free system on the ROS | Overlap Dice | ACU=97.5% | ✓ |
| 48-2018 | Thyroid nodule classification | Intra (264 benign-290 metastatic) Intra (5 benign-5 metastatic) inter (103 benign-103 metastatic) | 604 | Fusing VGG and class activation maps in one model | Sensitivity Specificity | TP/R=98% | ✓ |
| 49-2018 | Thyroid nodule classification | Local (850) thyroid nodule images 2,190 images from DDSI dataset | Not mentioned | ResNet-101 with dice loss, VGG-15, Renet/101/101 Classification trained on CNN with 101 algorithms | Sensitivity Specificity | 4. Accuracy=0.86, sensitivity=0.86, specificity=0.86 | ✓ |
| 50-2018 | Thyroid nodule classification | DDSI dataset | 298 | Data registration variation learning using pre-trained ResNet | Accuracy | Sensitivity Specificity | Accuracy=95% | ✓ |
| 51-2018 | Thyroid nodule classification | Local (164 images) DDSI dataset (202 images) | Not mentioned | Flow-based DeepLab model for lesion extraction Random Forest classifier | Dice coefficient and sensitivity, specificity, and BOC | Accuracy=0.86, sensitivity=0.86, specificity=0.86 and Dice coefficient=0.84 | ✓ |
| 52-2018 | Thyroid nodule classification | Local (2513) | 5082 | Thyroid nodules segmentation using CNNV5 and classification using CNN | ROC, AUC | AUC=94.51% | ✓ |
| 53-2018 | Thyroid nodule classification | Local (2213) | 624 | Thyroid nodules segmentation using CNN | TP/R, TP/R | Hausdorff distance | ✓ |
| 54-2018 | Thyroid nodule classification | Local (1857 thyroid nodule images) | Not mentioned | Classification of VGG features, Histogram of Oriented Gradient (HOG) and Local Binary Patterns (LBP) for classification | Accuracy | Sensitivity Specificity | Accuracy=98% | ✓ |

© The Institution of Engineering and Technology 2023
19 Prochazka, A., Gulati, S., Holinka, S. and Smutek, D.: ‘Patch-based classification of thyroid nodules in ultrasound images using direction independent features extracted by two-threshold binary decomposition’, Applied Soft Computing, 2019, 71, pp. 9–18

20 Yu, Q., Jiang, T., Zhou, A., Zhang, L., Zhang, C. and Xu, P.: ‘Computer-aided diagnosis of malignant or benign thyroid nodules based on ultrasound images’, European Archives of Oto-rhino-laryngology, 2017, 274, pp. 2891–2897

21 Acharya, U.R., Faust, O., Sree, S.V., Molinari, F. and Suri, J.S.: ‘ThyroScreen system: high resolution ultrasound thyroid image characterization into benign and malignant classes using novel combination of texture and discrete wavelet transform’, Computer methods and programs in biomedicine, 2012, 107, pp. 233–241

22 Tsantis, S., Dimitropoulos, N., Cavouras, D. and Nikiforidis, G.: ‘Morphological and wavelet features towards sonoagraphic thyroid nodules evaluation’, Computer-aided Medical Imaging and Graphics, 2009, 33, pp. 91–99

23 Ding, J., Cheng, H.D., Huang, J., Zhang, Y. and Liu, J.: ‘An improved quantitative measurement for thyroid cancer detection based on elastography’, European journal of radiology, 2012, 81, pp. 800–805

24 Acharya, U.R., Chowtiappa, P., Fujita, H., Bhat, S., Dua, S., Koh, J.E., Eugene, L.W.J., Kongmethbol, P. and Ng, K.H.: ‘Thyroid lesion classification in 242 patient population using Gabor transform features from high resolution ultrasound images’, Knowledge-Based Systems, 2016, 81, pp. 235–245

25 Chang, Y., Paul, A.K., Kim, N., Back, J.H., Chit, Y.J., Ha, E.J., Lee, K.D., Lee, H.S., Shin, D. and Kim, N.: ‘Computer/Aided diagnosis for classifying benign versus malignant thyroid nodules based on ultrasound images: A comparison with radiologist/Aided assessments’, Medical physics, 2016, 43, pp. 554–567

26 Chang, C.Y., Chen, S.J. and Tsai, M.F.: ‘Application of support-vector-machine-based method for feature selection and classification of thyroid nodules in ultrasound images’, Medical physics, 2010, 43, pp. 3494–3506

27 Koundal, D., Gupta, S. and Singh, S.: ‘Computer aided thyroid nodule detection system using medical ultrasound images’, Biomedical Signal Processing and Control, 2018, 40, pp. 117–130

28 Ouyang, F.S., Guo, B.L., Ouyang, L.Z., Liu, Z.W., Lin, S.J., Meng, W., Huang, X.Y., Chen, H.X., Qiu-Gen, H. and Yang, S.M.: ‘Comparison between linear and nonlinear machine-learning algorithms for the classification of thyroid nodules’, European journal of radiology, 2019, 113, pp. 251–257

29 Poudel, P., Illanes, A., Sheet, D. and Friebe, M.: Evaluation of commonly used nonlinear machine-learning algorithms for the classification of thyroid nodules on ultrasound images, Journal of healthcare engineering, 2018, 2018, pp. 251–257

30 AĞınAğık, AÇ., Abdelkadir, A., Lienkamp, S.S., Brox, T. and Ronneberger, O.: ‘3D U-Net: learning dense volumetric segmentation from sparse annotation’ Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, October 2016, pp. 424–432

31 AĞınAğık, AÇ., Abdelkadir, A., Lienkamp, S.S., Brox, T. and Ronneberger, O.: ‘Comparison of thyroid segmentation techniques for 3D ultrasound’, Proc. Medical Imaging 2017: Image Processing, Athens, Greece, February 2017, pp. 1013317

32 Zhou, S., Wu, H., Gong, J., Le, T., Wu, H., Chen, Q. and Xu, Z.: ‘Mark-Guided Segmentation of Ultrasonic Thyroid Nodules Using Deep Learning’ Proc. 2nd International Symposium on Image Computing and Digital Medicine, Chengdu, China, October 2018, pp. 21–26

33 Ronneberger, O., Fischer, P., Brox, T.: ‘U-net: Convolutional networks for biomedical image segmentation’ Proc. International Conference on Medical image computing and computer-assisted intervention, Munich, Germany, October 2015, pp. 234–241

34 Ma, J., Wu, F., Jiang, T.A., Zhu, J. and Kong, D.: ‘Cascade convolutional neural networks for automatic detection of thyroid nodules’, Medical physics, 2017, 44, pp. 1678–1691

35 Li, X., Wang, S., Wei, X., Zhu, J., Yu, R., Zhao, M., Yu, M., Liu, Z. and Liu, S.: ‘Fully Convolutional Networks for Ultrasound Image Segmentation of Thyroid Nodules’. Proc2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), Exeter, United Kingdom, June 2018, pp. 886–890

36 Ren, S., He, K., Girshick, R. and Sun, J.: ‘Faster r-cnn: Towards real-time object detection with region proposal networks’. Proc. Advances in neural information processing systems, Montreal, Quebec, Canada, December 2015, pp. 91–99

37 Li, H., Weng, J., Shi, Y., Gu, W., Mao, Y., Wang, Y., Liu, W. and Zhang, J.: ‘An improved deep learning approach for detection of thyroid papillary cancer in ultrasound images’, Journal of healthcare engineering, 2018, 8, pp. 1–12

38 Redmon, J. and Farhadi, A.: ‘YOLOv00900: better, faster, stronger’. Proc. IEEE conference on computer vision and pattern recognition, Honolulu, United States, July 2017, pp. 7263–7271

39 Szegedy, C., Ioffe, S., Vanhoucke, V. and Alemi, A.A.: ‘Inception-v4, inception-resnet and the impact of residual connections on learning’. Proc. Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, California, USA, February 2017, pp. 4278–4284

40 Wang, L., Yang, S., Yang, S., Zhao, C., Tian, G., Gao, Y., Chen, Y. and Lu, Y.: ‘Automatic thyroid nodule recognition and diagnosis in ultrasound imaging with the YOLOv2 neural network’, World journal of surgical oncology, 2019, 17, pp. 1–9

41 Li, X., Zhang, S., Zhang, Q., Wei, X., Pan, Y., Zhao, J., Xin, X., Qin, C., Wang, X., Liu, J. and Yang, F.: ‘Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonoigraphic images: a retrospective, multicohert, diagnostic study’, The Lancet Oncology, 2019, 20, pp. 193–201

42 He, K., Zhang, X., Ren, S. and Sun, J.: ‘Deep residual learning for image recognition’. Proc. IEEE conference on computer vision and pattern recognition, Las Vegas, USA, June 2016, pp. 770–778

43 ‘Darknet: Open Source Neural Networks in C’, [http://pjreddie.com/darknet/]

44 Song, J., Choi, Y.J., Masuoka, H., Park, S.W., Kim, S.J., Choi, J.Y., Kong, H.J., Lee, K.E., Lee, J., Kwak, N. and Yi, K.H.: ‘Ultrasound image analysis using deep learning algorithm for the diagnosis of thyroid nodules’, Medicine, 2019, 98, pp. e15133

45 Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z.: ‘Rethinking the inception architecture for computer vision’. Proc. IEEE conference on computer vision and pattern recognition, Las Vegas, USA, June 2016, pp. 2818–2826

46 Ko, S.Y., Lee, J.H., Yoon, J.H., Na, H., Hong, E., Han, K., Jung, I., Kim, E.K., Moon, H.J., Park, Y., and Lee, E.: ‘Deep convolutional neural network for the diagnosis of thyroid nodules on ultrasound’, Head & neck, 2019, 41, pp. 855–891

47 Ma, J., Wu, F., Jiang, T.A., Zhao, Q. and Kong, D.: ‘Ultrasound image-based thyroid nodule automatic segmentation using convolutional neural networks’, International journal of computer assisted radiology and surgery, 2017, 12, pp. 1895–1910

48 Nguyen, D.T., Pham, T.D., Batchuluun, G., Yoon, H.S. and Park, K.R.: ‘Artificial Intelligence-Based Thyroid Nodule Classification Using Information from Spatial and Frequency Domains’, Journal of clinical medicine, 2019, 8, pp. 1976

49 Song, W., Li, S., Liu, J., Qin, H., Zhang, B., Zhang, S. and Hao, A.: ‘Multitask cascade convolution neural networks for automatic thyroid nodule detection and recognition’, Journal of biomedical and health informatics, 2018, 23, pp. 1215–1224

50 Pedraza, L., Vargas, C., Narváez, F., Duráñ, O., Muñoz, E. and Romero, E.: ‘An open access thyroid ultrasound image database’. Proc. 10th International Symposium on Medical Information Processing and Analysis, Cartagena de Indias, Colombia, October 2015, pp. 92970W

51 Lee, J.H., Baek, J.H., Kim, J.H., Shim, W.H., Chung, S.R., Choi, Y.J. and Lee, J.H.: ‘Deep Learning-ÅBased Computer-Aided Diagnosis System for Localization and Diagnosis of Metastatic Lymph Nodes on Ultrasound: A Pilot Study’, Thyroid, 2018, 28, pp. 1322–1328

52 Wang, J., Li, S., Song, W., Qin, H., Zhang, B. and Hao, A.: ‘Learning from weakly-labeled clinical data for automatic thyroid nodule classification in ultrasound images’. Proc. 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, October 2018, pp. 3114–3118
53 Zhu, Y., Fu, Z. and Fei, J.: ’An image augmentation method using convolutional network for thyroid nodule classification by transfer learning’. Proc. 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, December 2017, pp. 1819–1823

54 Chi, J., Walia, E., Babyn, P., Wang, J., Groot, G. and Eramian, M.: ’Thyroid nodule classification in ultrasound images by fine-tuning deep convolutional neural network’. Journal of digital imaging, 2017, 30, pp. 477–486

55 Liu, T., Xie, S., Yu, J., Niu, L. and Sun, W.: ’Classification of thyroid nodules in ultrasound images using deep model based transfer learning and hybrid features’. Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, Louisiana, United States, March 2017, pp. 919–923

56 Bibicu, D., Moraru, L. and Biswas, A.: ’Thyroid nodule recognition based on feature selection and pixel classification methods’, Journal of digital imaging, 2013, 26, pp. 119–128
| Citation | Study Purpose | Dataset | Number of patients | 2D or 3D | Method | Measures | Results or main findings |
|----------|---------------|---------|--------------------|---------|--------|----------|-------------------------|
| [14]-2019 | Thyroid segmentation | 1.Local (151 US slices) 2.Optimac (168) | 2:4 healthy 2:2D | 2:2D | Watson tomosynthesis algorithm | Dice coefficient | 1. Acc=86.85% 2.DC=86.85% |
| [19]-2019 | Nodules classification | Local (90 thyroid nodules) | 60:2D | Two-Threshold Binary Decomposition + Random Forests (RF) and Support Vector Machine (SVM) classifiers | Accuracy Sensitivity Specificity | AUC-ROC SpecSens Specificity Sensitivity | 1. Accuracy=95%, Sensitivity=86.5%, Specificity=96%, AUC-ROC=0.87 for RF 2. Accuracy=85.5%, Sensitivity=95%, Specificity=86.9%, AUC-ROC=0.865 for SVM |
| [31]-2019 | Automated malignancy detection | Local (2012) | 4765:2D | Clinical features: size, cytopathology, margin, internal composition, shape, aspect ratio, vascularity, capsule, hyperechoic halo, cortical thickness, calcifications, real-time elastography grade, Nodules classification (Random forest, support vector machines, Adalboost, neural network, K-nearest neighbors, naive Bayes, support vector machines, convolutional neural network) | AUC Sensitivity Specificity | AUC-ROC=0.88 SpecSens=86.5% Specificity=85.3% Accuracy=85.7% |
| [35]-2018 | Thyroid, breast and external elastic laminae segmentation | 1. Optimac 2. 2011 MICCAI workshop | 1:1.6 2:2D | A combination of iterative random walks solver, random forest learning model and gradient vector flow based intramural belief propagation technique | Jaccard index Percentage of Area Difference | Handel Distance | 1.Jaccard index=0.96 2.Jaccard index=0.97 |
| [27]-2018 | Thyroid nodule detection | 1.Local 2.89 images from BI-RAD dataset | 1:150 2:Not mentioned | Speckle reduction using Neuroscopic Nukagam Total Variation, small rectangular region extraction and nodules segmentation using Neuroscopic Distance Regularized Level Set | True Positive False Positive True Positive False Positive True Positive False Positive True Positive False Positive | Dice Coefficient Overlap Metric Hausdorff Distance | Dice Coefficient=0.825 True Positive=0.794 False Positive=0.794 Overlap Metric=0.388 Hausdorff Distance=0.57 |
| [25]-2017 | Thyroid cancer detection | Local (260 nodules) | 187:2D | Sonographic features Reliability features selection+ Classification methods: extreme learning machine (ELM), SVM and fuzzy neural network | AUC-ROC Sensitivity Specificity | AUC-ROC SpecSens Specificity sensitivity | sensitivity=0.680.80 specificity=0.8788 positive predictive value=0.95 negative predictive value=0.9988 Youden index=0.9782 accuracy=0.9895 |
| [20]-2017 | Thyroid cancer detection | Local (100 images) | 543:2D | Region-based active contours morphological and texture features Classification: ANN and SVM | Dice coefficient Jaccard index Positive Predictive Value Sensitivity Specificity | Dice metric Jaccard index Specificity sensitivity | Dice metric=0.928 Jaccard index=0.928 sensitivity=0.909 specificity=0.923 positive predictive value=0.939 negative predictive value=0.923 Youden index=0.923 accuracy=0.923 |
| [33]-2017 | Thyroid cancer detection | Local (308) | Not mentioned:2D | Sonographic features: texture, contrast enhancement, hyperechoic halo, Nodule Echoic Features, Internal composition, Spatiotemporal, Calculations: Random Forest, Support Vector Machine and Logistic Regression | AUC-ROC Sensitivity Specificity | AUC-ROC SpecSens Specificity sensitivity | AUC-ROC=0.88 Sensitivity=85% Specificity=75% |
| [24]-2016 | Thyroid cancer detection | Local (242 thyroid US images) | 223:2D | 1. Image pre-processing using CLAHE 2. Extracting Feature: Shannon, Renyi, HOS, Kuper and Vejig entropy features from Gabor transformed coefficients 3. Feature space reduction and data normalization using locality sensitive discriminant analysis (LSDA) and Chilidze discriminant 4. Feature ranking using Relief-F 5. Classifications using SVM, ANN and MLP | Accuracy | Accuracy=84.3% |
| [25]-2016 | Thyroid cancer detection | Local (118 US images) | 59:2D | Feature set: histogram, intensity differences, elliptical fit, gray-level co-occurrence matrices and gray-level run-length matrices 3. Classifier: SVM | AUC-ROC | AUC-ROC=0.88 SpecSens=86% AUC-ROC=0.98 |
| [36]-2016 | Thyroid nodules segmentation | Local (240 thyroid US image of 42 subjects) | 42:2D | Spatial Neuroscopic Distance Regularized Level Set method based on Neuroscopic L-Mean clustering | Dice coefficient True Positive False Positive True Positive False Positive True Positive False Positive True Positive False Positive | Dice coefficient Overlap Metric Hausdorff Distance | Dice coefficient=0.825 true positive rate=0.985 false positive rate=0.772 mean absolute distance=1.8 Hausdorff Distance=0.76 Dice coefficient=0.825 |
| [37]-2015 | Multi-organ segmentation of thyroid, carotid artery, muscles and muscles | 1.Local (14 US images in B mode) 2. Local 18 US images in B mode from another hospital | 1:12 2:Not mentioned | Using spectral-related peaks to quantify the input US image | Dice coefficient | Dice coefficient=0.94 0.81, 0.83, 0.86, 0.88 2 DCC for thyroid, carotid, muscles and muscles 0.81, 0.83, 0.86, 0.87 |
| [56]-2013 | Thyroid nodule detection | 40 | 40:2D | Mean intensity has been used as the texture feature R-derivative statistical features evaluation using test and test | Accuracy Sensitivity Specificity | Accuracy=85% Sensitivity=90% Specificity=75% |
| [23]-2013 | Thyroid cancer detection based on Elastography | Local (225 color elastograms) | 125:Elastograms | 1. Lesion region extraction by radiologists 2. Statistical and texture features extraction 3. Feature selection and SVM classification | Sensitivity Specificity Accuracy | Sensitivity=84.6% Specificity=82.8% Accuracy=85.6% |
| [21]-2012 | Thyroid cancer detection | Local (400 benign and 400 malignant cases) | 20:1D | A combination of SWT (Discrete Wavelet Transform) and texture-based features (gray level co-occurrence matrices + Adalboost classifier) | TN TP FN TP | Acc Sensitivity Specificity and AUC | AUC-ROC=1 Acc sensitivity and specificity=100% |
| [30]-2010 | Thyroid nodule classifications | Local (1638 ROI) | 61:2D | Extracting 7 features from ROI—Two-SVMs for 4 types of nodules | Accuracy | Accuracy=100% |
| [18]-2008 | Thyroid segmentation | Local (20 US images) | 20:2D | 1. Locating Probable Thyroid Region 2. Adapting Weighted Median Filter 3. Remove the redundancy by Morphological Operation 4. Feature extraction using Haar Wavelet, Slantlet transform, BIPF Features and NNMD Features 5. Training SVM neural network to classify each block into thyroid gland and non-thyroid gland | Accuracy Sensitivity Specificity PPV NPV | Accuracy=0.86 PPV=0.62 Sensitivity=0.78 Specificity=0.87 PPV=0.68 NPV=0.84 |
| [22]-2008 | Multinodularity risk evaluation of thyroid nodules | Local (81) | 85:2D | A combination of morphological and texture features + SVM and Probabilistic neural networks (PNNs) | AUC-ROC Sensitivity Specificity and Likelihood ratio | AUC-ROC=0.84 Sensitivity=0.96 Specificity=0.83 Likelihood ratio=23.2 |