

Clinical Application of an Artificial Intelligence System for Diagnosing Thyroid Disease Based on a Computer Neural Network Deep Learning Model

Zhihai Li¹, Meilin Yin², Wenfeng Li³

¹Department of Ultrasound, Dalang Hospital, Dongguan, Guangdong, People's Republic of China; ²Department of Science and Education, Dalang Hospital, Dongguan, Guangdong, People's Republic of China; ³Department of Intervention, Dalang Hospital, Dongguan, Guangdong, People's Republic of China

Correspondence: Zhihai Li, Department of Ultrasound, Dalang Hospital, No. 85 Jinlang Middle Road, Dalang Town, Dongguan, Guangdong, 523777, People's Republic of China, Tel +86 18928276466, Email zhileidr123@outlook.com

Purpose: This study aimed to establish a stereoscopic neural learning network through deep learning and construct an artificial intelligence (AI) diagnosis system for the prediction of benign and malignant thyroid diseases, as well as repeatedly verified the diagnosis system and adjusted the data, in order to develop a type of AI-assisted thyroid diagnosis software with a low false negative rate and high sensitivity for clinical practice.

Patients and Methods: From July 2020 to April 2023, A total of 36 patients with thyroid nodules in our hospital were selected for diagnosis of thyroid nodules based on the *Expert Consensus on Thyroid Ultrasound*; samples were taken by aspiration biopsy or surgically and sent for pathological diagnosis. The ultrasonic diagnosis results were compared with the pathological results, a database was established based on the ultrasonic diagnostic characteristics and was entered in the AI-assisted diagnosis software for judgment of benign and malignant conditions. The data in the software were corrected based on the conformity rate and the reasons for misjudgment, and the corrected software was used to evaluate the benign and malignant conditions of the 36 patients, until the conformity rate exceeded 90%.

Results: The initial conformity rate of the AI software for identifying benign and malignant conditions was 88%, while that of the software utilizing the database was 94%.

Conclusion: We established a stereoscopic neural learning network and construct an AI diagnosis system for the prediction of benign and malignant thyroid diseases, with a low false negative rate and high sensitivity for clinical practice.

Keywords: artificial intelligence-aided diagnosis software, computer neural network deep learning model, thyroid nodules

Introduction

The thyroid is a crucial endocrine organ in the human body. Thyroid cancer, one of the most prevalent endocrine system malignancies, comprises the largest proportion among head and neck malignancies.¹ It consists of two connected lobes, and is one of the largest endocrine glands in the human body, weighing 20–30 g in adults. Thyroid lesions are often found on the gland, with a prevalence of 4–7%. Most of them are asymptomatic, and thyroid hormone secretion is normal.² The American Thyroid Association (ATA) recommends ultrasonic inspection as the method of choice for preoperative diagnosis of thyroid nodules due to its efficiency, non-invasiveness, non-radioactivity, high repeatability, and low cost in evaluating benign and malignant thyroid nodules. The percentage of incidental malignancy is 18.42%.³ However, this method has limitations in terms of clinical diagnosis. Coupled with the rapid growth in the number of patients, the limited medical resources are stretched, increasing the workload of sonologists and reducing the average time spent on diagnosis, which further affects the precision of ultrasonic diagnosis. As a result, unnecessary aspiration biopsies and even

diagnostic procedures are often conducted in clinical work. Following excessive or unnecessary diagnostic procedures, patients are required to take thyroid hormones for an extended period of time,² and the adjustment of drug dose and repeated monitoring of tumors may result in deteriorating the patients physical, mental, and spiritual health. In addition, the high incidence of postoperative complications³ and relapse⁴ may have a negative impact on the mental health of patients, particularly young and middle-aged women, resulting in a surge in doctor-patient disputes and huge medical compensation in China.

Traditional studies on artificial intelligence (AI) systems rely primarily on two methods. First, a recognition model is designed using classical pattern recognition. In such studies, doctors must collaborate closely with engineering researchers, who must comprehend all aspects of cognitive objectives and then manually model them. However, cognitive objectives may appear differently on medical images, and a single model cannot adequately cover all images. In most cases, long-term and repeated trials are conducted to develop complex models. Second, the deep neural networks (DNN) learning model is used for “End to End” training.⁵ This is a method for automatically discovering features that enables adding of areas or images to be trained to the input images rather than defining the model, thereby minimizing the loss of input and verified data.

In recent years, DNN learning has emerged as a prominent research method, and is also used in this study. This mode can reduce the problem to its simplest form and automatically construct a complex model. For example, some researchers use shift learning models trained using digital camera images⁶ (eg, YOLO, ImageNet); however, medical images are different from ordinary digital camera images and image formation principles, hence creating an original model of shift learning using medical images will be optimal. Due to the differences in various types of medical images (eg, CT, MRI), target sites and imaging methods, it is necessary to develop shift learning models for each case in this stage. Our project team designed a system that combines preprocessing, network composition, and loss function for medical images to support the R&D of the thyroid AI system.

Materials and Methods

General Data

Participants: A total of 36 patients with thyroid disease who were diagnosed and treated in our hospital (Dalang Hospital, Dongguan City, Guangdong Province, China) between July 2020 to April 2023 (this group included: 1. patients with thyroid nodules; 2. patients from whom samples were taken by aspiration biopsy or surgically and sent for pathological diagnosis) were selected. The selection criteria of patients are selected from the conventional patient group. When the thyroid nodules are found in the routine examination, patients with high malignant degree will be screened for inclusion in the study according to the relevant guidelines for thyroid diagnosis and treatment in China, and patients will be communicated with, sign the informed consent for interventional puncture biopsy, and then sent for pathological examination after surgery.

Instruments and equipment: A color Doppler ultrasound machine, computer (attached with the AI identification system and image acquisition system: PACS medical system and DI-com image acquisition system).

This study was conducted in accordance with the declaration of Helsinki and approved by the Ethics Committee of Dalang Hospital. Written informed consent was obtained from all participants.

Methods

The Standard Procedure for Diagnosing Thyroid Disease Was Established

1. Measurement of transverse diameter of left and right lobes, anteroposterior diameter, and anteroposterior diameter of isthmus:

Standard section: The transverse section revealed the planes of common carotid artery and internal jugular vein.

Transverse diameter: The horizontal distance between the lateral lobe envelopes of the thyroid was measured through the anterior wall of the common carotid artery.

Anteroposterior diameter: The vertical distance between the anterior and posterior thyroid envelopes was measured through the lateral wall of the trachea.

Anteroposterior diameter of isthmus: The vertical distance between the anterior and posterior envelopes was measured in the center of the isthmus.

2. Measurement of the superior-inferior diameter of left and right lobes:

Standard section: The vertical section revealed the vertical distance between the tip and basal envelopes measured on the entire lateral lobe section:

- a. Normal value for thyroid in adults: No unified standard measurement.
- b. The superior-inferior diameter of left and right lobes of thyroid was 3~6.0 cm, the bilateral diameter was 2~3.0 cm, and the anteroposterior diameter was 1~2.0 cm, of which the anteroposterior diameter was significant and was generally not more than 2 cm.
- c. Sequence of diameter lines: Long diameter → bilateral diameter → anteroposterior diameter.
- d. Shrinkage of thyroid: Bilateral and anteroposterior diameters <1.0 cm (based on thyroid echo and experimental tests).

3. Measurement principles:

No measurement was required when the size was normal and there was no specific clinical requirement.

Measurement was required in the following cases

- ① Enlargement of bilateral lobes of thyroid.
- ② Enlargement of unilateral lobe of thyroid.
- ③ Shrinking of thyroid.
- ④ After partial thyroidectomy.

4. Suspected malignant nodules: a. Three diameter lines were given in the following order: long diameter — lateral diameter — anteroposterior diameter; the distance between the nodule and the upper pole, lower pole, anterior capsule, and posterior capsule as well as its relationship to the capsule were measured (blurred, interrupted, unclear boundary with adjacent anterior cervical muscle).

5. Specification for stock images:

- a. 2D ultrasonography: Transverse section of the center of thyroid; vertical and transverse sections of the lateral lobes; vertical and transverse sections of the lesion site; and vertical and transverse sections of the lymph nodes under suspicion.
- b. Color Doppler of lesions and lymph nodes.

(Note: in the actual work, information of the acquired images is identified by human eyes; however, this method is affected by many factors, such as myopia, partial color deficiency, degree of fatigue, and perception of the disease, with high negative and false positive rates.)

6. Routine diagnostic features of malignant nodules:

- (1) Nodule boundary: Malignant nodules have irregular boundaries, with burr and convex papillate.
- (2) Aspect ratio of nodules: An aspect ratio greater than 1 indicates an elevated risk of malignancy.
- (3) Calcification: Diffuse grainy calcification in a hypoechoic solid tumor mass or more scattered grainy calcification in papillary and umbrellalike parenchymal echo appear to be manifestations of malignant nodules in a mixed tumor mass.
- (4) Intraparenchymal blood vessels: The presence of irregular halos around the nodule signifies a high likelihood of malignancy.
- (5) Low echo: The presence of a low echo in a thyroid nodule indicates an increased risk of malignancy.

- (6) Abundant blood flow in nodules: Blood flow is abundant in thyroid nodules, and current research indicates that these nodules are more likely to harbor malignancies.

The physician made the initial diagnosis for patients with thyroid disease and collected the corresponding images and image features of thyroid diseases; samples taken by aspiration biopsy or surgically were sent for pathological diagnosis (the diagnosis was made based on the thyroid-related hormone tests for some very typical cases, such as hyperthyroidism). If the diagnosis results were consistent with the pathological results, image features of this case were determined, if not, ultrasonic images of the misdiagnosed case were located, and the case was discussed to determine the missing image features. A database was compiled based on the image characteristics of all cases in preparation for the development of the AI diagnosis system. The examiners doing ultrasound examination are basically colleagues in the ultrasound department of our hospital, all of whom have the title of treating physician or above, have rich work experience, and have received corresponding training in the implementation of the project operation. Attention should be paid to the standardization of the image storage of the selected cases during the training inspection.

Construction of a stereoscopic deep learning neural network model.

Deep learning networks are configured in two stages: division and clustering networks. As demonstrated by the FAST CNN network (enhanced convolutional neural network) and YOLO, designing the network for each purpose can increase the accuracy of each task, thereby enhancing the overall performance with minimal data.⁷

Our procedure begins with the training of the division and clustering networks. The output of the division network is in pixel precision, whereas the output of the clustering network is the establishment of each category. This pre-learning, similar to a weak identifier, can be used to obtain coarse characteristics.

The two networks were then embedded into the special loss function and trained as a single network. The incorporation of the output of the division network into the clustering network facilitated the selection of the objective's features within the clustering network. In addition, the loss function could access two networks, making it more efficient at learning than a single network and capable of resolving under-learning and over-learning issues.

Establishment of AI Diagnosis System

The grayscale interval was set dynamically based on the grayscale value of the learned image. The original image was sliced in the grayscale interval set. After slicing, the histogram of images in each luminance interval was computed and scaled to the same level. As the input of the CNN network, the obtained histogram of the index in each luminance interval was used to construct the characteristic model. This simulated the thought process of a doctor making a diagnosis, without being constrained by the form in which the disease was transferred onto the perspective image. Furthermore, after amplification at each luminance level, it was easier to learn the local characteristics of luminance level.

The stereoscopic neural network deep learning model consisted of pre-processing (feature spatial transformation, feature abstraction), establishing the stereoscopic identification model, and establishing the multiphase prediction model.

According to the luminance value interval of the target lesions, such as tumor mass and calcification, the original data were segmented and the luminance value was amplified proportionally, enabling us to capture the target features in the local luminance interval to the maximum extent. When time permitted, we also adjusted the coordinates of the center point, retaining only the meaningful data areas. Pre-processing yielded a set of data that was unaffected by the size and shape of the target lesion area.

The stereoscopic identification model was used to identify features such as tumor mass and calcification, and to provide a comprehensive lesion index for the thyroid lesions observed currently.⁸ First, an independent neural network was created for each luminance group to identify and define the features in each luminance value interval. Using the independent neural networks of each luminance as the input layer, we defined the deep network of the template as the middle layer node, and constructed the stereoscopic deep network that was used to identify the model. The stereoscopic deep network can effectively identify complex targets with multiple features by integrating and learning multidimensional features. It can also effectively identify various forms of target lesions based on pre-processing.

The DTT (digital thyroid tomosynthesis) multiphase prediction model learns the normal thyroid (negative) morphology using the multiphase imaging data of DTT, thereby forming the standard thyroid characteristic groups in all directions. The impact of the original data in all directions was defined as a sequence of images, with each image

divided into several small regions. We defined only a portion of the valid image as the learning subject, matched the small areas in all directions, and classified them as a sequence of images. By comparing the unknown image to the prediction model, we could obtain the matching degree in each small area. There may be a high likelihood of lesions in a region with a low degree of matching. The DTT multiphase prediction model can not only help in accurately calculating the lesion area, but also detect unknown lesions that were not discovered in the learned data, thereby reducing the false positive and false negative rates significantly.

The AI diagnosis system is developed by Qingdao Hanlin Huili Technology Co., LTD. The system uses deep learning methods to build a fusion ultrasound image, identify different modal characteristics, coordinate and integrate multiple deep learning networks into a three-dimensional neural network, establish a comprehensive thyroid artificial intelligence diagnosis system, and verify the diagnosis system based on retrospective and prospective case data, with a view to providing a low false negative rate for the clinic. Highly sensitive thyroid AI-assisted diagnostic aid.

Interaction Between Intelligent Diagnosis System and Daily Work and Data Enhancement

An AI aided diagnosis system was established with data from the thyroid disease database using related software, and the data was fed back to clinical examination to improve the diagnosis efficiency and indirectly boost economic benefits. At the same time, related data were collected to enhance the content and scope of the database.

The physician evaluated the stock images of previous patients with thyroid disease using the AI diagnosis system. The manifestation of no nodules by the AI system is depicted in Figure 1. If nodules were diagnosed, they were divided into the benign group (Figure 2) and malignant group (Figure 3). The conformity rates of ultrasonic diagnosis and pathological results were compared. If AI group > routine diagnosis group, the AI diagnosis system had clinical significance; if AI group < routine diagnosis group, the AI diagnosis system should be enhanced in the following way: If the diagnosis results of the AI diagnosis system were consistent with the pathological results, the image features of this case were determined; if not, ultrasonic images of the misdiagnosed case were found, and the case was discussed to find the missed image features; subsequently, data were added to the database to enhance the AI diagnosis system (a repetitive process).

Statistical Processing

Data in this study were analyzed using SPSS 18.0 statistical software. Measurement data are expressed as mean \pm standard deviation, and were compared using the *t*-test; enumeration data are expressed as percentage, and were compared using the chi-squared test. $P < 0.05$ indicated statistical differences.



Figure 1 Manifestation of no nodules by AI diagnosis. For the first measurement of the isthmus of the thyroid, the symbol is “+”; for the second measurement, the left and right diameters of the right lobe are “++”, the anteroposterior diameters of the right lobe are “x”, the left and right diameters of the left lobe are “□”, and the anteroposterior diameters of the left lobe are “Δ”.

Table 1 Comparison of Diagnosis Results Between Routine Diagnosis Group and Pathological Group

Case Group \ Category	Thyroid Cancer			Nodular Goiter			Adenoma	
	Papillary	Follicular	Medullary	With Cystic Change and Bleeding	With Calcification	With Cystic Change, Bleeding, and Calcification	Follicular	Papillary
Routine diagnosis group (n)	6	0	0	8	5	4	9	4
Pathological group (n)	9	1	1	6	4	3	9	3

Table 2 Comparative Analysis of Benign and Malignant Outcomes in the Routine Diagnosis, Pathological, and AI Groups

Benign Group \ And Malignant	PTC	Benign
Routine diagnosis group (n)	6	30
Pathological group (n)	11	25
AI group (n)	14	22

Table 3 Comparison of Benign and Malignant Diagnosis Outcomes Based on a Modified AI Database and Pathological Group

Benign Group \ And Malignant	PTC	Benign
Pathological group (n)	11	25
AI group (n)	12	24

Discussion

“The Bethesda System for Reporting Thyroid Cytopathology” (TBSRTC)⁹ includes six diagnostic categories, each with an implied risk of malignancy (ROM) and recommended clinical management. Bethesda I: Nondiagnostic or Unsatisfactory; Bethesda II: Benign; Bethesda III, Atypia of Undetermined Significance (AUS) or Follicular Lesion of Undetermined Significance (FLUS); Bethesda IV, Follicular Neoplasm (FN) or Suspicious for a Follicular Neoplasm (SFN); Bethesda V, Suspicious for Malignancy (SFM); Bethesda VI, Malignant. Thyroid nodules that fall within Bethesda categories III–IV have an overall risk of malignancy of between 15% and 40%. Incidental malignancy was found in 1.53% (8/522) cases of Bethesda II.¹⁰ The rate of malignancy is 18.1% for Bethesda category III (AUS/FLUS) patients with MNG and 20.1% for those with SNG.³

In this study, we performed routine diagnosis on 36 patients for the first time, and then sent the samples for pathological diagnosis. The results revealed that the diagnostic accuracy of the routine physician group was approximately 72%, and the conformity rate of benign and malignant cancer was approximately 83%. Image characteristics of the routine diagnosis results consistent with the pathological diagnosis results were extracted to establish a database, and AI diagnosis software was configured to evaluate the benign and malignant conditions of the 36 patients. The conformity of the pathological results was approximately 88%, which was higher than the conformity rate of the routine physician

group but lower than we anticipated (90%). The low conformity rate is due to the following factors. 1. Not all acquired images were standardized; 2. Not all cases were diagnosed by the same physician, which increased the diagnostic error rate; 3. Different performance of the color Doppler ultrasound machine decreased the diagnostic conformity rate; 4. Due to economic considerations, aspiration biopsy or surgery was frequently applicable to a single or multiple targets of interest; however, the neglected targets could be missed and affect the conformity rate. We extracted image features based on experience and added them to the AI database, resulting in a 94% conformance rate, which met expectations. Reasons for this difference are as follows: Images acquired were not standard, and some of them were also incomplete, leading to a low recognition degree.

Based on these results, the AI image diagnosis technology has advanced by leaps and bounds along with the development of computer technology, electronic engineering, statistics, and other fields. The combination of artificial neural networks and deep learning systems with medical research has produced fruitful results, particularly the computer-assisted diagnosis system (CAD). In the last few years, technological developments in the surgical field have been rapid and are continuously evolving. One of the most revolutionizing breakthroughs was the introduction of the IoT concept within the surgical practice.¹¹ Since 2008, when Lim et al proposed the CAD system based on the thyroid nodule ultrasonic images,¹² numerous studies have reported a significant enhancement in the ability of CAD combined with AI to distinguish benign from malignant thyroid nodules. Zhang et al constructed an ANN diagnostic model and confirmed that its accuracy, sensitivity, and specificity were 90%, 88.24%, and 90.91%, respectively.¹³ Kriegeskorte et al identified benign and malignant thyroid nodules in 162 patients using DNN and ultrasonic images, with 96.34% accuracy, 86% sensitivity, and 99% specificity, respectively.¹⁴ However, none of these methods were evaluated based on clinical practice, and some complex images required complex calculational analysis and lengthy processing times, so these methods are not suitable to clinical practice. The accuracy, sensitivity, and specificity of the CAD system for real-time monitoring were 84.6%, 80%, and 88.1%.¹⁵ However, regions of interest must be artificially marked, and values must be assigned based on ultrasonic characteristics to distinguish benign from malignant nodules; consequently, they remain inapplicable to clinical practice in China.

In this study, based on the development and clinical application of the computer neural network deep learning model for the thyroid disease AI diagnosis system, several deep learning networks were integrated into a single stereoscopic neural network to create a comprehensive thyroid AI diagnosis system. A total of 36 patients with thyroid disease were diagnosed twice, and their samples were collected via aspiration biopsy or surgically and sent for pathological diagnosis. Data obtained from each link were processed, a database was established and entered into the computer-aided diagnosis system, and the accuracy and efficiency of the computer-aided diagnosis system were verified repeatedly using this method; The final results demonstrated that the accuracy and efficiency of the CAD system in the second diagnosis were significantly higher than in the group receiving routine diagnosis.

In addition, this method was validated using retrospective and prospective data, which can be conducted in an almost infinitely repeatable manner, indicating that it has great diagnostic potential for clinical diagnosis.

The rapid development of computer technology makes introspection on its applications essential. AI technology could eventually develop to the point where its accuracy, efficiency, breadth, depth, strength, and adaptability outstrip those of humans. This could be a boon to humanity or a catastrophe. New moral standards and legal norms are necessary for the advancement of AI technology.

Conclusions

We established a stereoscopic neural learning network and construct an AI diagnosis system for the prediction of benign and malignant thyroid diseases, with a low false negative rate and high sensitivity for clinical practice.

Funding

This study was supported by Key Project of Dongguan Social Development Science and Technology in 2022 (NO.: 2040912860690).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Alexander EK, Cibas ES. Diagnosis of thyroid nodules. *Lancet Diabetes Endocrinol.* 2022;10(7):533–539. doi:10.1016/S2213-8587(22)00101-2
2. Mulita F, Anjum F. *Thyroid Adenoma*. Treasure Island (FL): StatPearls Publishing; 2023.
3. Mulita F, Plachouri MK, Liolis E, Vailas M, Panagopoulos K, Maroulis I. Patient outcomes following surgical management of thyroid nodules classified as Bethesda category III (AUS/FLUS). *Endokrynol Pol.* 2021;72(2):143–144. doi:10.5603/EP.a2021.0018
4. Hazkani I, Edwards E, Stein E, et al. Perioperative outcomes in children with Hashimoto's thyroiditis undergoing total thyroidectomy. *Am J Otolaryngol.* 2023;44(2):103785. doi:10.1016/j.amjoto.2022.103785
5. Li X, Li X, Xie C, et al. cGAS guards against chromosome end-to-end fusions during mitosis and facilitates replicative senescence. *Protein Cell.* 2022;13(1):47–64. doi:10.1007/s13238-021-00879-y
6. Satrya WF, Yun JH. Combining Model-Agnostic Meta-Learning and Transfer Learning for Regression. *Sensors.* 2023;23(2):583. doi:10.3390/s23020583
7. Sabri ZS, Li Z. Low-cost intelligent surveillance system based on fast CNN. *PeerJ Comput Sci.* 2021;7:e402.
8. Korenaga TK, Tewari KS. Gynecologic cancer in pregnancy. *Gynecol Oncol.* 2020;157(3):799–809. doi:10.1016/j.ygyno.2020.03.015
9. Ali SZ, Cibas ES. *The Bethesda System for Reporting Thyroid Cytopathology: Definitions, Criteria and Explanatory Notes*. New York: Springer; 2010.
10. Mulita F, Iliopoulos F, Tsilivigkos C, et al. Cancer rate of Bethesda category II thyroid nodules. *Med Glas.* 2022;19(1). doi:10.17392/1413-21
11. Mulita F, Verras GI, Anagnostopoulos CN, Kotis K. A Smarter Health through the Internet of Surgical Things. *Sensors.* 2022;22(12):4577. doi:10.3390/s22124577
12. Lim R, Ellett LK, Roughead EE, Cheah PY, Masnoon N. Patient-Reported Questionnaires to Identify Adverse Drug Reactions: a Systematic Review. *Int J Environ Res Public Health.* 2022;19(18):11209. doi:10.3390/ijerph191811209
13. Zhang J, Zhang X, Sh Y, Liu B, Hu Z. Diagnostic AI Modeling and Pseudo Time Series Profiling of AD and PD Based on Individualized Serum Proteome Data. *Front Bioinform.* 2021;1:764497. doi:10.3389/fbinf.2021.764497
14. Kriegeskorte N, Golan T. Neural network models and deep learning. *Curr Biol.* 2019;29(7):R231–R236. doi:10.1016/j.cub.2019.02.034
15. Le EPV, Wang Y, Huang Y, Hickman S, Gilbert FJ. Artificial intelligence in breast imaging. *Clin Radiol.* 2019;74(5):357–366. doi:10.1016/j.crad.2019.02.006

Journal of Multidisciplinary Healthcare

Dovepress

Publish your work in this journal

The Journal of Multidisciplinary Healthcare is an international, peer-reviewed open-access journal that aims to represent and publish research in healthcare areas delivered by practitioners of different disciplines. This includes studies and reviews conducted by multidisciplinary teams as well as research which evaluates the results or conduct of such teams or healthcare processes in general. The journal covers a very wide range of areas and welcomes submissions from practitioners at all levels, from all over the world. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-inflammation-research-journal>