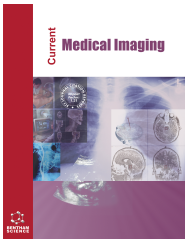
BENTHAM SCIENCE

CrossMark

Current Medical Imaging

Content list available at: https://benthamscience.com/journals/cmir



RESEARCH ARTICLE

Super-resolution based Nodule Localization in Thyroid Ultrasound Images through Deep Learning

Jing Li[1], Qiang Guo[1,*], Shiyi Peng[1] and Xingli Tan[1]

[1]*Department of Ultrasound, Qingpu Branch of Zhongshan Hospital Affiliated to Fudan University, No. 1158 Gongyuan East Road, Qingpu District, Shanghai, 201700, China*

Abstract:

Background:

Currently, it is difficult to find a solution to the inverse inappropriate problem, which involves restoring a high-resolution image from a low-resolution image contained within a single image. In nature photography, one can capture a wide variety of objects and textures, each with its own characteristics, most notably the high-frequency component. These qualities can be distinguished from each other by looking at the pictures.

Objective:

The goal is to develop an automated approach to identify thyroid nodules on ultrasound images. The aim of this research is to accurately differentiate thyroid nodules using Deep Learning Technique and to evaluate the effectiveness of different localization techniques.

Methods:

The method used in this research is to reconstruct a single super-resolution image based on segmentation and classification. The poor-quality ultrasound image is divided into several parts, and the best applicable classification is chosen for each component. Pairs of high- and low-resolution images belonging to the same class are found and used to figure out which image is high-resolution for each segment. Deep learning technology, specifically the Adam classifier, is used to identify carcinoid tumors within thyroid nodules. Measures, such as localization accuracy, sensitivity, specificity, dice loss, ROC, and area under the curve (AUC), are used to evaluate the effectiveness of the techniques.

Results:

The results of the proposed method are superior, both statistically and qualitatively, compared to other methods that are considered one of the latest and best technologies. The developed automated approach shows promising results in accurately identifying thyroid nodules on ultrasound images.

Conclusion:

The research demonstrates the development of an automated approach to identify thyroid nodules within ultrasound images using super-resolution single-image reconstruction and deep learning technology. The results indicate that the proposed method is superior to the latest and best techniques in terms of accuracy and quality. This research contributes to the advancement of medical imaging and holds the potential to improve the diagnosis and treatment of thyroid nodules.

Keywords: Super-resolution, Thyroid nodule, Classification, Accuracy, Validation, Training, Nodule localization.

Article History	Received: December 13, 2023	Revised: February 18, 2024	Accepted: February 26, 2024

1. INTRODUCTION

Thyroid nodules, including goitre, adenoma, and malignant growths, are quite common in the clinic. However, the patient may get the wrong therapy at the wrong time if the preoperative diagnosis is inaccurate. In the previous four decades, the number of cases of thyroid nodule lesions in the United States has grown by up to 3 percent annually [1 - 3]. Ultrasound imaging has several benefits, such as cheap cost, real-time imaging, high sensitivity, no biopsy required, no physical harm, and repeatability. Ultrasound may be used to detect thyroid nodules and the surrounding tissues, providing valuable information for distinguishing between benign and malignant conditions.

The implementation of The Bethesda System for Reporting

* Address correspondence to this author at the Department of Ultrasound, Qingpu Branch of Zhongshan Hospital Affiliated to Fudan University, No. 1158 Gongyuan East Road, Qingpu District, Shanghai, 201700, China;
E-mail: 1d0_uxgs8a6nu2@qphospital.com

Thyroid Cytopathology (TBSRTC) in its 2023 version has greatly simplified the procedure of reporting thyroid fine needle aspirations. It classifies findings into six separate categories, each associated with a unique risk of malignancy (ROM). The classifications are as follows: (i) nondiagnostic, (ii) benign, (iii) atypia of unknown significance (AUS), (iv) follicular neoplasm, (v) suggestive for malignancy, and (vi) malignant. This method offers a consistent reporting structure and enhances the ROM for each category using the latest data. It delivers an average ROM and a range of cancer risk assumptions. The AUS category is subdivided into two subcategories based on the ROM and molecular profile, which improves the accuracy of the diagnosis. In addition, TBSRTC incorporates the most recent advancements in molecular testing and adjusts its nomenclature to match the 2022 WHO Classification of Thyroid Neoplasms. This guarantees that the classification remains up-to-date with the changing medical standards and practices [4].

In addition, the Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) provides a systematic method for categorising thyroid fine needle aspiration (FNA) cytology findings. Category II (Benign) has a malignancy risk of around 4% (2–7%), recommending conservative treatment. Category III (Atypia of Undetermined Significance/Follicular Lesion of Undetermined Significance, AUS/FLUS) has a higher complexity in risk, with studies showing a malignancy rate of up to 25% for patients who have had two consecutive AUS/FLUS diagnoses and undergo surgery. The results emphasise the crucial significance of making follow-up and management decisions based on FNAC results within these categories [5].

The blurriness of thyroid nodules, poor resolution, excessive noise in the photographs, the biological complexity of the thyroid, and the degree of skill of the radiologist all contribute to the difficulties of ultrasound imaging [6]. In this study, deep learning for lesion recognition has been used as a means of localization and diagnosis in thyroid ultrasonography. Ultrasound imaging, in contrast to invasive methods like MRI and CT scans, is painless and easy. Ultrasound imaging occurs in real-time and is continuous [7]. In ultrasonic imaging, the detection depth is related to the method's sensitivity. Acoustic shadowing and post-echo amplification may have an effect on the results. These characteristics are used for the diagnostic ultrasound pictures. A large number of training instances are required to guarantee that supervised classification accuracy is not compromised by the curse of dimensionality from which it suffers. Training resources that can be put to use in the real world are often few. If training samples are few, semi-supervised learning, in which some unlabeled examples are included in the training set in accordance with prescribed rules, may be able to help. Several semi-supervised pixel-level classification algorithms have also been developed in recent studies [6, 7]. Deep learning methods have been extensively used for the detection and diagnosis of thyroid nodules in ultrasound pictures based on the findings of the search. Reconstructing high-quality pictures from low-resolution ones is the goal of these approaches, which may increase the accuracy as well as the efficiency with which nodule localization is performed. In general, the use of deep learning strategies has a significant promise for the precise and time-saving identification of thyroid nodules in ultrasound images. However, it is essential to guarantee that these models are interpretable in order to provide physicians with the capacity to comprehend and have faith in the judgements that are made by algorithms. In addition, more research is required to verify these methods using bigger datasets and in a variety of clinical contexts. Despite the fact that there is very little research available on the subject, it was found that the problem of insufficient training samples also has a substantial influence on the accuracy of the super-resolution mapping. First, we need to consider the spectral variation present in training samples of the same class. The generated super-resolution variation and collaborative representation are then used to determine the abundance percentages of each class inside the ultrasound image.

The thyroid gland, a significant endocrine organ, has a normal weight range of 20-30 grams in adults. It is common to find lesions on the gland, with a prevalence rate of 4-7%. However, these lesions are mostly asymptomatic and do not typically affect the secretion of thyroid hormones. This information could be integrated into the existing literature and further detailed insights could be obtained from articles on PubMed [8].

1.1. Literature Review

Dukker and coworkers developed a technique for identifying lesions using breast ultrasonography in 2002 [9]. The Bayesian classifier made a correct diagnosis once the RGI established the lesion severity. In 2007, Eystratios *et al.* [10] proposed an automated nodule diagnostic method using thyroid ultrasound characteristics. Thyroid borders were determined for diagnostic purposes using the TBD method and then categorised using the K-means algorithm. In their study, Huang *et al.* [11] proposed a technique that employed segmented breast ultrasound images. Using a combination of RGB and particle swarm optimisation, this method was able to get either perfect or almost perfect settings. Shan *et al.* [12] in 2014 endorsed ASPS as a means to boost regional development. The histogram of the ultrasound image and the geographic parameters of the breast lesion were used to pinpoint the exact placement of the seed in this technique. The bulk of the findings were from breast ultrasounds. Ultrasound imaging of the thyroid, which has a more complex echo structure than that of the breast, provides better diagnostic information. These factors complicate the application of machine learning for diagnostic categorization of thyroid medical ultrasonography pictures. Based on the results, a system is developed to identify and categorise thyroid conditions from ultrasound images. In this method, lesions are classified using Adam's classifier and then located with an improved FCN trained with transfer learning. First, the deep learning strategy is used to locate and label suspicious spots on a thyroid ultrasound, with promising results. Ultrasonography is often employed in the screening procedure for thyroid cancer because of its inexpensive cost. Nodules may be classified as solid or cystic based on ultrasonographic criteria related to their pathophysiology [13, 14]. Ultrasound imaging may occasionally differentiate normal thyroid gland tissue apart from thyroid nodules, which is useful

for parameter estimation. However, ultrasound images may be difficult to separate anatomically due to their granular speckle pattern and lack of contrast between bodily sections. Ultrasound may be used in a number of different ways to divide up thyroid nodules [10, 14 - 22]. Most segmentation approaches use human-drawn "seed boundaries" to begin with. This seed is used to initiate the algorithm. To estimate the limits of a B-mode picture, a user may generate a "seeded boundary" by manually drawing a border of nodules. The generation of a seed precludes the use of an algorithm in real-time. Therefore, verification must be carried out retrospectively. Clinical segmentation in real-time may become achievable *via* the segmentation of seedless thyroid nodules. Feature hierarchies are constructed using visual models like convolutional networks. Convolutional networks that are trained from the ground up and pixel-by-pixel outperform existing methods in semantic segmentation. In a study [23], "fully convolutional" networks were constructed, which, *via* fast inference and learning, may output a size proportionate to any input. The strategy [23] is referred to as "Method 1." The fully convolutional networks may be used for spatially dense prediction challenges and combined with existing models. Here, AlexNet, VGG Net, and Google Net are convolutionalized and then their representations are adjusted to enable them to do image segmentation. In another study [24], the effectiveness of three non-automatic segmentation methods was evaluated and contrasted for free-form 3D ultrasound imaging. 'Method 2' will refer to the procedure outlined in a previous study [24]. Precision, robustness, user-friendliness, required human participation, and processing time are only a few of the metrics that have been used to evaluate these algorithms and draw comparisons. Furthermore, another technique was proposed, which was referred to as "Method 3" [25], for differentiating between thyroid nodules. In this method, expanded convolutional layers are used in a revolutionary multi-output convolutional neural network approach. Method 3 distinguishes between the normal thyroid gland, thyroid nodules, and the cystic components inside the nodules. Nodules on the thyroid gland may be solid, somewhat cystic, or mostly cystic. Thyroid nodules may be broken down into cystic components to reveal their internal anatomy. Ultrasound scans show hypoechoic cysts. It is vital to distinguish cysts, which may include hypoechoic patches, from the arteries and veins of the thyroid gland, which may also seem hypoechoic on ultrasound. Thyroid capillaries are responsible for secreting this. The segmentation process begins with the ultrasound imaging of the thyroid gland and then checks the nodules for cystic components. More anatomical characteristics may be used for ultrasonic image classification as the graphics processing units can handle bigger and more complex neural networks. In this regard, methods from deep learning may be helpful [26, 27]. Even without a seed, deep learning algorithms may be operated in real time because of their fast inference time (measured in milliseconds). In order to differentiate between normal thyroid glands and thyroid nodules using B-mode pictures, this study introduces a multi-pronged convolutional neural network or CNN. The programme can detect and analyse several thyroid anatomical features in real-time. Thyroid nodules may be detected, evaluated, and mapped using this technology. The approach is verified by applying it to a manually segmented mask, and the results are compared to those obtained using the standard seeding method.

2. METHODS

2.1. Image Segmentation

It is well known that earlier techniques for image super-resolution reconstruction included the use of a large training database. This database consisted of millions of image pairings with varying degrees of resolution. Learning from such a massive database requires a lot of computer power and memory, so it is not an option for everyone. In addition, the input picture has a complex structure and background, similar to ultrasound images. Due to this, some of the high-frequency information in the reconstructed high-resolution picture would be lost, which is not a desired outcome. An experiment of super-resolution reconstruction on an ultrasound image is executed. This is necessary because of the substantial variety of features that may be found in various areas. The reconstructed high-resolution images are provided, each of which is trained with a unique set of image categories. When an ultrasound image is provided with a variety of different textures, it is segmented into many different texture areas. Then, a super-resolution reconstruction of each segment is performed using the dictionary learned using the category database that corresponds to it. In this research, the input low-resolution picture is segmented using the multi-scale normalised cut technique [28], and the outcome of the resolution enhancement is presented in Fig. (**1**).

2.2. Classification of Thyroid Nodule Sections

As the input image is split up into multiple distinct parts, classifying each segment into the category is the most appropriate step. This approach, which improves picture super-resolution reconstruction by processing each section of the input image with its own lexicon, is very valuable. SPM has become more popular in image classification [28 - 30], which has seen a great deal of success in many applications. The picture is segmented into spatial subregions that are progressively finer, and histograms of the image's local characteristics are computed for each of these subregions. Each individual section of the low-resolution picture is provided as its own image. After creating the "descriptor" layer, the initial step is to extract a set of scale-invariant feature transform (SIFT) descriptors (*P - dimensional*) from the ultrasound image feature point. These descriptors are represented by the equation $D = [d_1, d_2,....,d_N] \in R^{PxN}$. After that, an *M*-entry codebook called $C = [c_1, c_2,....,c_N] \in R^{PxM}$ turns each descriptor into an M-dimensional code. As a consequence of this, each code is considered to be an individual component of the "code" layer, and these codes together constitute the image. The local correlations between similar descriptors to improve segment categorization need to be considered, which requires similar codes. After that, the codes are computed using the following formula:

$$\min_{x} \sum_{i=1}^{N} \|d_i - Cx_i\|^2 + \lambda \|s_i \odot x_i\|^2 \quad s.t. \quad x_i = 1, \forall i \qquad \textbf{(1)}$$

Where $X = [x_1, x_2,....,x_N] \in R^{MxN}$ computes similarity to the input descriptor d_j in the following way: where element-wise multiplication is denoted by $\odot$, $s_i \in R^M$ is similar to the d_j such that:

$$si = exp\ \ [\![((dist(d\_i, C)))/\sigma]\!] \qquad (2)$$

Here, $dist(d_i, C)=[dist(d_i, c_1),...,dist(d_i,c_M)]^T$ denotes Euclidean distance between d_j,C_j. Here, σ is a weight parameter for s_i. After that, $max[dist(d_i, C)]$ is used to transform s_i into the interval (0,1). In practise, the resultant code x_i only contains a few significant values; thus, other minuscule coefficients are zero. Moreover, the other tiny coefficients are also zero. The "SPM" layer histogram is formed by averaging and normalising several codes from each sub-region. "Max" pooling is utilised when the "max" function is applied row-wise to produce a vector x_{in} of the same size as its input. After that, the pooled features are normalised using the L2-norm function as $\_out=x\_in/||x\_in||\_2$. After that, each histogram is appended to the others to produce a vector. When it comes to tagging objects in an image, this vector serves as the most accurate depiction. The input picture has a poor resolution; hence, this method is used to determine the category for each individual part of the image.

3. SUPER-RESOLUTION OF INDIVIDUAL SEGMENTS

Following the step of assigning each section of the input low-resolution picture to the category that best fits it, each section's high-resolution picture is reconstructed using the thyroid ultrasound database [31]. The image super-resolution reconstruction is a low-resolution scene reconstruction problem. Furthermore, a low-resolution picture (X) is transformed into a high-resolution one (Y) of the same scenario. A high-resolution picture is input, Y = DBX, where D is the down-sampling operator, B is the blurring operator, and X is an image with an unknowingly higher resolution improved in the previous step. As every image X satisfies the potential to be the answer to this inverse problem, it is abundantly clear that this is a very poorly framed issue.

3.1. An Incomplete Sparse Representation

Suppose $x \in R^n$ represents the patches that were taken from the resolution-improved image X, and that $x \in R^m$ ill represents the patches that were taken from the high-resolution images Y, then the high-resolution vector is represented by V_{High} for high-resolution image patches. The formula for this combination is as follows $x = D_H\sigma$,

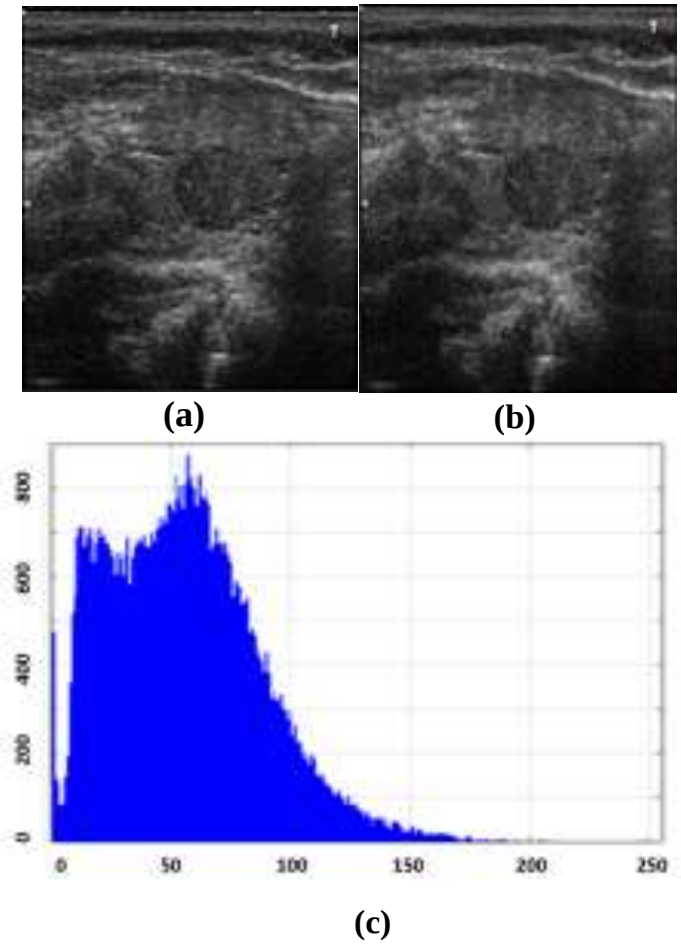


Fig. (1). Qualitative results (**a**) input image (**b**) resolution enhanced image (**c**) histogram of resolution enhanced image.

Where $V_H \epsilon R_{nxK}$ ($n<K$) represents a vector that corresponds to the high-resolution vector and α ε $R^K(\|V_0\|<<K)$ is coefficients that contain few non-zero values. After that, the sparse coefficient a is computed by representing the image patches y that have low-resolution with the low-resolution pixel vector V_{low} contained with the high-resolution pixels V_{high}. Sparse coefficients may describe low-resolution patches $y = V_{low}\alpha$. Then, given a low-resolution patch y, we will calculate the sparse coefficients in terms of C_{low} and utilise them with the high-resolution pixels vector V_{high} to construct the high-resolution patch y. The formula for calculating the sparse coefficients α is as follows:

$$\min_{\alpha}\|\alpha\|_0, s.t. \|PV_{low}\alpha - Py\|_2^2 \leq \varepsilon \quad \textbf{(3a)}$$

$$\max_{\alpha}\|\alpha\|_0, s.t. \left\|PV_{high}\alpha - Py\right\|_2^2 \geq \varepsilon \quad \textbf{(3b)}$$

L0-norm($\|.\|_0$) and (P) are linear feature extraction operators that should remove the key parts of the low-resolution picture as customers prefer high-frequency material.

As the feature extraction operator, a high-pass filter of some form is often used as the appropriate option. In this article, the features are extracted using the following four 1 - D filters:

$$p_1 = [-1,0,1], p_2 = p_1^T \quad \textbf{(4a)}$$

$$p_3 = [1,0,-2,0,1], p_4 = p_3^T \quad \textbf{(4b)}$$

As a consequence of this, four descriptor feature vectors will be constructed for each patch by using the aforementioned four filters. As the final representation of the low-resolution patch, a vector is created by concatenating all of the aforementioned elements together. Afterward, L1-norm is minimised into L0-norm,

$$\min_{\alpha}\|PV_{low}\alpha - Py\|_2^2 + \lambda\|\alpha\|_1 \quad \textbf{(5a)}$$

$$\max_{\alpha}\left\|PV_{high}\alpha + Py\right\|_2^2 + \lambda\|\alpha\|_1 \quad \textbf{(5b)}$$

Where λ balances coefficient sparsity and approximation accuracy for y. After the ideal solution, α, has been found, the high-resolution patch may be reconstructed using the equation $x = V_{high}\alpha^*$ As a consequence of this, the high-resolution picture Y is produced by combining all of the reconstructed high-resolution patches and then taking an average of the region where they overlap.

3.2. Generation of Training Data using High-resolution Image

Deep learning often depends on the mining of big datasets that are representative of their subject matter. This allows for the training of a resilient network that generalises effectively when it is put into action in the real world. Using ultrasound data from a database [31] for training is a basic process. Now, the dataset includes each associated low-resolution ultrasound image and super-resolved image as a foundation for a diversified training dataset that incorporates a number of different variants. In a high-resolution image, target patches consist of several microbubbles of varying intensities, each of which includes multiple target patches. In this article, the assumption that thyroid nodule is detected by using singular value filtering or contrast-enhanced imaging sequences has been used.

3.3. The Structure of Deep Neural Network

Deep neural networks employ three-layer-block encoding contraction. Each block has two leaky rectified linear units (ReLU)-triggered 3x3 convolution layers and a 2x2 Max-pooling operation. All convolution layers employ leaky ReLUs [32] to avoid idle neurons or nodes affecting model performance. Batch normalisation before network activation improves trainability. Moreover, hyper-parameter optimisation has some limitations, but it allows quicker learning [33]. A dropout layer randomly disables 50% of the following latent layer. The decoder creates a comprehensive map from concealed information. A 2x2 nearest-neighbor up-sampling layer repeats picture rows and columns, and two 5x5 deconvolution layers, one with an output stride of 2, make up the decoder [23]. The final block has two deconvolution layers, the second of which preserves its output stride value of 2. A five-by-five convolution with a linear activation function converts feature space to a single-channel picture.

3.4. The Optimizer

To reduce the cost function, the network is trained on batches of 256 imaging frames over 20,000 iterations using the Adam optimizer with a learning rate of 0.001. This cost function matches [34] and is represented as:

$$c(x, Y|\theta) = \| f(x|\theta) - G_\sigma * Y\|_2^2 + \lambda\|f(Y|\theta)\|_1 \quad \textbf{(6)}$$

The input patches are denoted by x, while the target super-resolved patches are denoted by super-resolved image Y. The nonlinear neural network function is denoted by $f(\theta)$, and its parameters (weights and biases) θ are shown in brackets. The regularisation parameter λ supports sparse network predictions and is conservatively set at 0.01. The G_σ standard deviation is initialised at 1 pixel upon training. The suggested technique highlights thyroid nodules during image reconstruction. Each cycle generates new data online, which improves the model's robustness and capacity to generalise. Dropout while training was found to help in this regard. The encoded latent space characteristics are randomly deactivated with a 0.5 probability.

4. RESULTS AND DISCUSSION

This work employed 173 ultra-sonographic pictures of 173 thyroid samples (80 normal and 93 cancerous) from Kaggle [31]. The Intel i7-6900K Nvidia GeForce GTX 1080Ti system was used to run MATLAB2021a. Fig. (**1a-c**) shows pre-processing inputs and output images. Fig. (**2**) shows localization results, which were achieved by all proposed methods after super-resolution segmentation (Figs. **2a**,**b**) and optimization (Figs. **2c**,**d**). The thyroid tissues share morphology with the lesion site, but current methods cannot find them. The proposed methods can localize the nodule with border

infiltration, lesions with ambiguous boundaries, and generally clear nodules, as shown by the boundary box and the coloured area in Fig. (**2**). In the proposed work, the neural network training requires ground truth. Thyroid nodule detection is the best technique, which uses intersection over union (IoU), Dice loss, accuracy, sensitivity, specificity, ROC, and AUC in order to assess the disease diagnosis. Fig. (**3**) shows the ROC analysis of the proposed method and state of art methods. The proposed method's specificity value increases when ROC approaches 1. Fig. (**4a,b**) plots training and validation performance: (a) Dice loss and (b) Accuracy. Table **1** presents the performance evaluation of the proposed method before super-resolution segmentation and after super-resolution segmentation techniques. The proposed method outperforms other approaches by using a deep neural network average IoU of 0.81 for 173 test samples. The findings are poor since the photographs are classified without considering their substance. Our approach analyses image textures to distinguish healthy from damaged tissues. Our technique is more flexible and accurate. The proposed method will leverage the classifier for further categorization processes. This improves method comparison. The proposed method's ROC integrates various methodologies, and the curve is steeper than other methods. Table **2** compares the proposed method with the recent methods [23 - 25]. As indicated in the table, our technique has significantly improved classification accuracy.

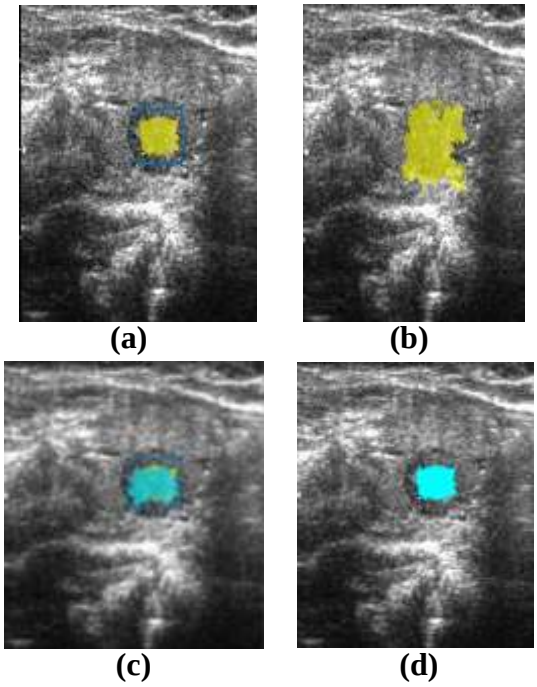


Fig. (2). Qualitative results (**a**) nodule localization with boundary after super-resolution segmentation (**b**) region detection after super-resolution segmentation (**c**) Nodule localization with boundary after optimization (**d**) region detection after optimization.

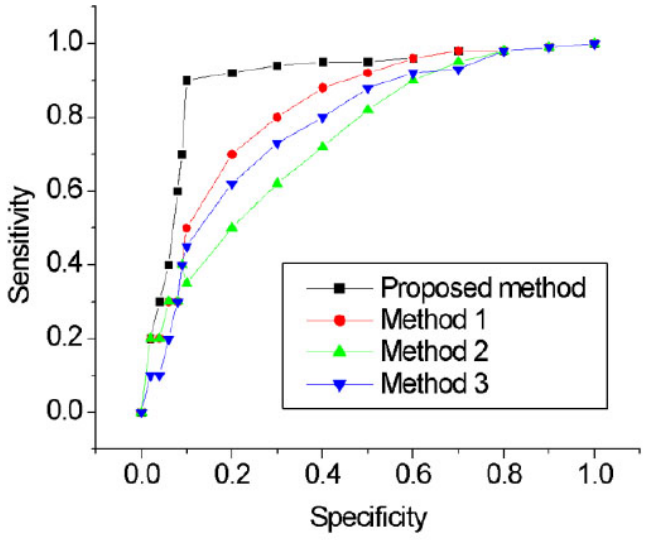


Fig. (3). ROC analysis of the proposed method and state-of-the-art methods.

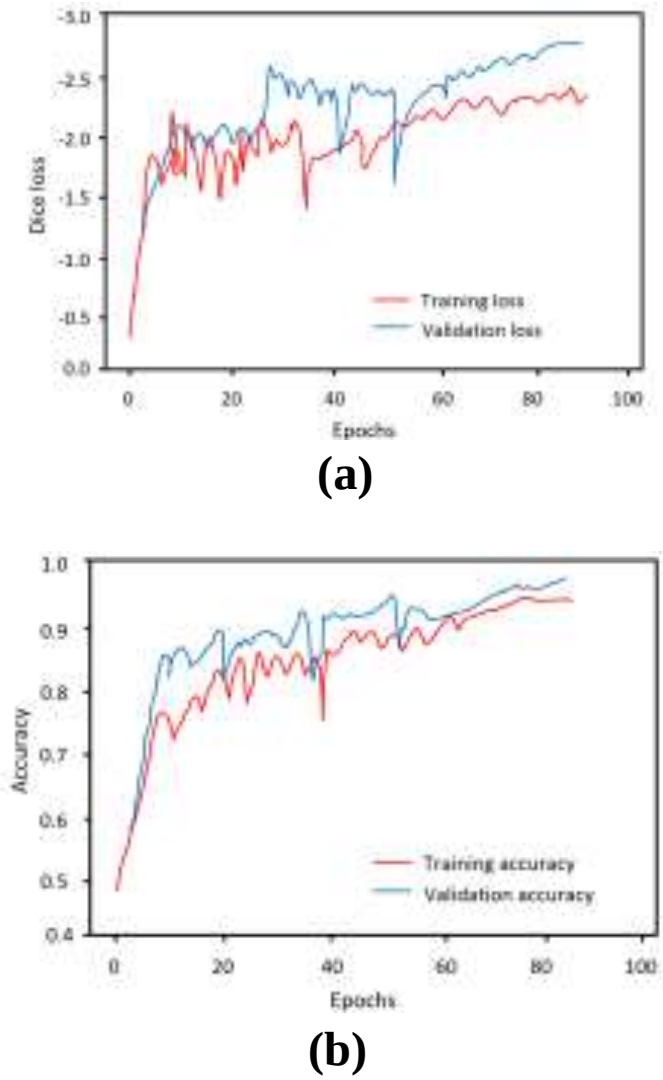


Fig. (4). Training and validation performance (**a**) dice loss (**b**) accuracy.

Table 1. Performance evaluation of the proposed method.

Method	Accuracy	AUC	Mean IoU
Before super-resolution segmentation	0.72	0.81	0.53
After super-resolution segmentation	0.81	0.83	0.71
After classification	0.83	0.87	0.78
After optimization	0.91	0.96	0.81

Table 2. Comparison of the proposed method with recent state-of-the-art methods.

Method/Refs.	Method Used	Performance Metric
Method 1 [23]	Fully convolutional neural networks	0.622 (mean IoU)
Method 2 [24]	Segmentation and improvement using machine learning	0.862 and 0.876 (Dice coefficient)
Method 3 [25]	Multi-output convolutional neural network algorithm	0.76 (Dice coefficient)
Proposed method	Super-Resolution based nodule localization in thyroid ultrasound images	0.81

Recent research by Mulita *et al.* (2022) compared early postoperative complications between total thyroidectomy (TT)

and subtotal thyroidectomy (STT) for differentiated thyroid cancer. Their findings indicated no significant difference in complications, such as hematoma, wound infection, hypoparathyroidism, and temporary recurrent laryngeal nerve palsy between the two surgical approaches. This information highlights the importance of not only improving diagnostic and imaging techniques, as our method does but also considering the surgical strategies employed in the management of thyroid conditions. While our approach focuses on enhancing the accuracy of thyroid lesion detection, it is crucial to recognize that the ultimate goal is to optimize patient outcomes. The choice between TT and STT, as demonstrated by Mulita *et al.* (2022), plays a vital role in minimizing postoperative complications and ensuring the well-being of patients [35].

Localization errors may affect benign or malignant disease diagnosis. Segmented ROIs resemble lesion areas. The segmented ROI is too big and contains useless lesion data. The first algorithm will be limited by picture ROI. A wide region covers the ROI. Thus, non-lesion tissue data is abundant. Thyroid ultrasound pictures are difficult to interpret, increasing the chance of location-related false positives. Shadows may be misinterpreted for the ROI. If a lesion-suspected site is linked to the image border, the whole lesion-containing region is recognised. Thyroid ultrasonography may distort the picture and leave a worthless area. This approach cannot ensure diagnostic accuracy. In this regard, lesion localization inside the ROI can be improved, and visual interference can reduce sample adaptability. The proposed approach uses deep learning and classification so that the proposed model may learn about image features and transmit segmentation abilities *via* training. Thus, it adapts better to complex thyroid lesion tissue.

CONCLUSION AND FUTURE WORK

This article suggests automated thyroid lesion detection. The thyroid ultrasound image is segmented using the proposed deep learning-based method to find the lesion. Next, the Adam classifier is used for the localization region. However, further testing is needed to establish whether the lesion is malignant. Contract trials for this study employed many methods on the same data set. The experimental findings imply that the proposed method-based lesion localization technology may obtain an IoU value of 0.81, outperforming earlier methods. This approach can diagnose benign and malignant thyroid nodules with accuracy, AUC, and IOU values of 0.91, 0.96, and 0.81. These experiments prove the approach works. Even compared to older approaches, the new method exceeds the state-of-the-art in detecting lesion sites and discriminating benign from malignant thyroid lesions. Lesion detection methods may be improved. GAN networks may recreate thyroid ultrasound pictures. Higher-resolution images will improve thyroid nodule auto-identification. Moreover, the use of deep learning to the problem of pinpointing nodules in thyroid ultrasound pictures may provide a few different ways. Researchers may investigate multi-modality fusion, where ultrasound pictures are fused with additional medical images like CT or MRI to enhance nodule localisation in ultrasound images. The term “transfer learning” refers to the practise of bringing learned skills from one situation to bear on another similar one. When working with limited data, researchers have the option of using transfer learning to create more precise nodule localisation models. Despite their outstanding performance on a variety of medical imaging tasks, deep learning models are sometimes referred to as “black boxes” due to the difficulty in understanding how they arrived at a given interpretation. To improve the model's credibility with healthcare practitioners and to foster mutual understanding, researchers might investigate explainable deep learning models. The majority of deep learning models used for nodule localisation in thyroid ultrasound images are currently intended for post-hoc analysis, meaning they cannot be used for real-time diagnosis. During ultrasound procedures, doctors may benefit from real-time diagnostic systems. However, researchers need to evaluate these procedures further.

LIST OF ABBREVIATIONS

AUC	=	Area under the curve
TBSRTC	=	Thyroid Cytopathology
ROM	=	Risk of Malignancy
FNA	=	Fine Needle Aspiration
SIFT	=	Scale-invariant Feature Transform
TT	=	Total Thyroidectomy
STT	=	Subtotal Thyroidectomy

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used that are the basis of this study.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The data used to support the findings of this study are included in the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Bibbins-Domingo K. Screening for thyroid cancer: US preventive services task force recommendation statement. JAMA 2017; 317(18): 1882-7.

[2] Moon W-J. Benign and malignant thyroid nodules: US differentiation multicenter retrospective study. Radiology 2008; 247(3): 762-70. [http://dx.doi.org/10.1148/radiol.2473070944]

[3] Frates MC. Management of thyroid nodules detected at US: Society of radiologists in ultrasound consensus conference statement Radiology 2005; 237(3): 794-800. [http://dx.doi.org/10.1148/radiol.2373050220]

[4] Ali SZ, Baloch ZW, Cochand-Priollet B, Schmitt FC, Vielh P,

VanderLaan PA. The 2023 Bethesda System for Reporting Thyroid Cytopathology. Thyroid 2023; 33(9)
[http://dx.doi.org/10.1089/thy.2023.0141] [PMID: 37427847]

[5] Zahid A, Shafiq W, Nasir KS, *et al.* Malignancy rates in thyroid nodules classified as Bethesda categories III and IV; A subcontinent perspective. J Clin Transl Endocrinol 2021; 23: 100250.
[http://dx.doi.org/10.1016/j.jcte.2021.100250] [PMID: 33643850]

[6] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015; 521(7553): 436.
[http://dx.doi.org/10.1038/nature14539]

[7] Brito JP, Gionfriddo MR, Al Nofal A, *et al.* The accuracy of thyroid nodule ultrasound to predict thyroid cancer: Systematic review and metaanalysis. J Clin Endocrinol Metabol 2014; 99(4): 1253-63.
[http://dx.doi.org/10.1210/jc.2013-2928]

[8] Mulita F, Anjum F. Thyroid Adenoma. In: StatPearls. Treasure Island (FL): StatPearls Publishing 2023.
[PMID: 32965923]

[9] Drukker Karen. Computerized lesion detection on breast ultrasound. Medical Physics 2002; 29(7): 1438.
[http://dx.doi.org/10.1118/1.1485995]

[10] Keramidas EG. Efficient and Effective Ultrasound Image Analysis Scheme for Thyroid Nodule Detection. In: Image Analysis and Recognition ICIAR 2007 Lecture Notes in Computer Science. Berlin, Heidelberg: Springer 2007.
[http://dx.doi.org/10.1007/978-3-540-74260-9_93]

[11] Huang Q. Optimized graph-based segmentation for ultrasound images. Neurocomputing 2014; 129.SI: 216-24.
[http://dx.doi.org/10.1016/j.neucom.2013.09.038]

[12] Shan J. A novel automatic seed point selection algorithm for breast ultrasound images. 2008 19th International Conference on Pattern Recognition,. Tampa, FL, 08-11 December 2008, pp. 1-4.

[13] Mazzaferri EL. Management of a solitary thyroid nodule New England J Med 1993; 328(8): 553-9.

[14] Chang C-Y, Lei Y-F, Tseng C-H, Shih S-R. Thyroid segmentation and volume estimation in ultrasound images. IEEE Trans Biomed Eng 2010; 57(6): 1348-57.
[http://dx.doi.org/10.1109/TBME.2010.2041003]

[15] Maroulis D E, Savelonas M A, Iakovidis D K, Karkanis S A, Dimitropoulos N. Variable background active contour model for computer-aided delineation of nodules in thyroid ultrasound images. IEEE Trans Inf Technol Biomed 2007; 11(5): 537-43.
[http://dx.doi.org/10.1109/TITB.2006.890018]

[16] Iakovidis D K, Savelonas M A, Karkanis S A, Maroulis D E. A genetically optimized level set approach to segmentation of thyroid ultrasound images. Appl Intell 2007; 27(3): 193-203.
[http://dx.doi.org/10.1007/s10489-007-0066-y]

[17] Singh N, Jindal A. A segmentation method and comparison of classification methods for thyroid ultrasound images. Int J Comput Appl 2012; 50(11): 43-9.

[18] Mahmood N H, Rusli A H. Segmentation and area measurement for thyroid ultrasound image. Int J Sci Eng Res 2011; 2(12): 1-8.

[19] Kollorz E N, Hahn D A, Linke R, Goecke T W, Hornegger J, Kuwert T. Quanti_cation of thyroid volume using 3-D ultrasound imaging. IEEE Trans Med Imag 2008; 27(4): 457-66.

[20] Nugroho HA, Nugroho A, Choridah L. Thyroid nodule segmentation using active contour bilateral filtering on ultrasound images. 2015 International Conference on Quality in Research (QiR),. Lombok, Indonesia, 10-13 August 2015, pp. 43-46.
[http://dx.doi.org/10.1109/QiR.2015.7374892]

[21] Tsantis S, Dimitropoulos N, Cavouras D, Nikiforidis G. A hybrid multi-scale model for thyroid nodule boundary detection on ultrasound images Comput Methods Programs Biomed 2006; 84(2-3): 86-98.
[http://dx.doi.org/10.1016/j.cmpb.2006.09.006]

[22] Selvathi D, Sharnitha VS. Thyroid classification and segmentation in ultrasound images using machine learning algorithms. 2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies,. Thuckalay, India, 21-22 July 2011, pp. 836-841.
[http://dx.doi.org/10.1109/ICSCCN.2011.6024666]

[23] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition,. UC Berkeley, USA, 2015, pp.3431-3440.

[24] Poudel P, Illanes A, Sheet D, Friebe M. Evaluation of commonly used algorithms for thyroid ultrasound images segmentation and improvement using machine learning approaches. J Healthc Eng 2018; 2018(Sep): 1-13.
[http://dx.doi.org/10.1155/2018/8087624] [PMID: 30344990]

[25] Kumar V, Webb J, Gregory A, *et al.* Automated segmentation of thyroid nodule, gland, and cystic components from ultrasound images using deep learning. IEEE Access 2020; 8: 63482-96.
[http://dx.doi.org/10.1109/ACCESS.2020.2982390] [PMID: 32995106]

[26] Kumar V, Webb JM, Gregory A, *et al.* Automated and real-time segmentation of suspicious breast masses using convolutional neural network. PLoS One 2018; 13(5): e0195816.
[http://dx.doi.org/10.1371/journal.pone.0195816] [PMID: 29768415]

[27] Looney P, Stevenson GN, Nicolaides KH, *et al.* Automatic 3D ultrasound segmentation of the first trimester placenta using deep learning. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017),. Melbourne, VIC, Australia, 18-21 April 2017, pp. 279-282.
[http://dx.doi.org/10.1109/ISBI.2017.7950519]

[28] Cour T, Benezit F, Shi 1. Spectral segmentation with multiscale graph decomposition. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05),. San Diego, CA, USA, 2005, pp. 1124-1131.

[29] Mosseri I, Zontak M, Irani M. Combining the power of Internal and External denoising. IEEE International Conference on Computational Photography (ICCP),. Cambridge, MA, USA, 2013, pp. 1-9.
[http://dx.doi.org/10.1109/ICCPhot.2013.6528298]

[30] Wang J, Yang J, Yu K, *et al.* Locality-constrained Linear Coding for image classification. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition,. San Francisco, CA, USA, 2010, pp. 3360-3367.

[31] DDTI: Thyroid Ultrasound Images. Available from: https://www.kaggle.com/datasets/dasmehdixtr/ddti-thyroid-ultrasound-images

[32] Xu B, Wang N, Chen T, Li M. Empirical evaluation of rectified activations in convolutional network. arXiv:150500853 2015; 2015: 00853.
[http://dx.doi.org/10.48550/arXiv.1505.00853]

[33] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv:150203167 2015; 2015: 03167.
[http://dx.doi.org/10.48550/arXiv.1502.03167]

[34] Nehme E, Weiss LE, Michaeli T, Shechtman Y. Deep-STORM: super-resolution single-molecule microscopy by deep learning. Optica 2018; 5(4): 458-64.
[http://dx.doi.org/10.1364/OPTICA.5.000458]

[35] Mulita F, Verras GI, Dafnomili VD, *et al.* Thyroidectomy for the management of differentiated thyroid carcinoma and their outcome on early postoperative complications: A 6-year single-centre retrospective study. Chirurgia 2022; 117(5): 556-62.
[http://dx.doi.org/10.21614/chirurgia.2736] [PMID: 36318685]