

# Automatika

Journal for Control, Measurement, Electronics, Computing and Communications



ISSN: (Print) (Online) Journal homepage: [www.tandfonline.com/journals/taut20](http://www.tandfonline.com/journals/taut20)

## Residual U-Net approach for thyroid nodule detection and classification from thyroid ultrasound images

Sheeja Agustin, Sruthy S, Ajay James & Philomina Simon

To cite this article: Sheeja Agustin, Sruthy S, Ajay James & Philomina Simon (2024) Residual U-Net approach for thyroid nodule detection and classification from thyroid ultrasound images, *Automatika*, 65:3, 726-737, DOI: [10.1080/00051144.2024.2316503](https://doi.org/10.1080/00051144.2024.2316503)

To link to this article: <https://doi.org/10.1080/00051144.2024.2316503>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 26 Feb 2024.



Submit your article to this journal [↗](#)



Article views: 945



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 4 View citing articles [↗](#)



# Residual U-Net approach for thyroid nodule detection and classification from thyroid ultrasound images

Sheeja Agustin<sup>a</sup>, Sruthy S<sup>b</sup>, Ajay James<sup>c</sup> and Philomina Simon<sup>d</sup>

<sup>a</sup>Department of Computer Science & Engineering, Marian Engineering College, Thiruvananthapuram, India; <sup>b</sup>The Department of Computer Science and Engineering, St. Joseph's College of Engineering and Technology, Palai, India; <sup>c</sup>Department of Computer Science and Engineering, Government Engineering College, Thrissur, India; <sup>d</sup>Department of Computer Science, University of Kerala, Thiruvananthapuram, India

## ABSTRACT

With so many thyroid knobs (nodules) discovered by accident, it is critical to recognize as many aberrant knobs (nodules) as possible from fine-needle aspiration (FNA) biopsies or other medical procedures while excluding those that are virtually certainly benign. Thyroid ultrasonography, on the other hand, is prone to interobserver variability and subjective translations. An effective deep learning model for segmenting and categorizing thyroid nodules in this study follows the stages below: data collection from a well-known archive, The Thyroid Digital Image Database (TDID), which comprises ultrasound pictures from 298 patients, preprocessing using anisotropic diffusion filter (ADF) for removing noise and enhancing the images, segmentation using a bilateral filter for segmenting images, feature extraction using grey level occurrence matrix (GLCM), feature selection using Multi-objective Particle Swarm with Random Forest Optimization (MbP-SRA) and finally classification happens were Residual U-Net will be used. Experiment evaluation states the proposed model outperforms well than other state-of-art models.

## ARTICLE HISTORY

Received 26 September 2023  
Accepted 10 December 2023

## KEYWORDS

Classification; deep learning; residual U-Net; segmentation; thyroid nodule; ultrasound images

## 1. Introduction

Thyroid nodule (TN) detection has grown dramatically during the last two decades, with many more nodules being discovered by chance. Because many thyroid nodules are benign or act in a nonthreatening manner, establishing whether they are benign or malignant via fine needle analysis (FNA) biopsies and/or surgery can save patients time and money. Sonography is commonly used to check these atypical thyroid nodules. Radiologists have identified hypoechogenicity, microcalcifications, hardness and a taller-than-wide morphology as sonographic indications of thyroid nodules that signal malignancy [1]. 'Not suspicious', 'probably benign', 'one suspicious feature', 'two suspicious features', 'three or more suspicious features' and 'probable malignancy' are the TI-RADS scores for 'not suspicious', 'probably benign', 'one suspicious feature', 'two suspicious features', 'three or more suspicious features' and 'probable malignancy', respectively [2]. On the other hand, TI-RADS TN assessment takes a long time and is rarely correct. Because current sonographic criteria for identifying malignant nodules are insufficient, and variability in thyroid nodule echo patterns restrict radiologists' judgement capability [3] radiologists' accuracy mainly relies on personal experience.

The deployment of a validated TI-RADS reporting mechanism would allow clinicians to stop doing routine thyroid ultrasound exams with the assurance that they would not miss a malignancy. It would also give doctors a suggested follow-up approach for patients with a moderate-risk TI-RADS score, sparing the system a lot of unneeded imaging tests and the stress that comes with them. Furthermore, because sonographic properties may be obtained in ultrasound pictures that can be digitized and fed into a machine learning system, radiologists can use TI-RADS scores to learn how the computer classifies the image [4–8]. Thyroid nodules are groups of thyroid cells that have developed improperly in the thyroid gland. The thyroid is an endocrine gland that produces and disperses hormones throughout the body. Thyroid hormones are produced and released by the thyroid gland, which is made up of two lobes connected by an isthmus (or 'bridge'). The isthmus is 1.2 cm long and 1.2 cm broad, and each lobe is pear-shaped and 5 cm long and 2.5 cm wide. It regulates thyroid hormone secretion, which is important for body temperature regulation, and has a big impact on children's intellect and growth. It also creates hormones that aid in the body's metabolic management. Thyroid hormone production that is excessive

**CONTACT** Sheeja Agustin ✉ [sheejaagustin.cse@gmx.com](mailto:sheejaagustin.cse@gmx.com) Department of Computer Science & Engineering, Marian Engineering College, 695582, Thiruvananthapuram

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

or insufficient (due to a thyroid that is too big or too tiny, respectively) produces pathological alterations and thyroid abnormalities. As a result, the volume of the thyroid gland is frequently used to identify abnormal thyroid symptoms. Thyroid illnesses are divided into three categories: hyperthyroidism, hypothyroidism, and Hashimoto's thyroiditis.

Thyroid hormones regulate a number of activities in the body, including body temperature, digestion and heart rate [8]. These hormones play a vital role in protein synthesis, body temperature regulation, and total energy generation and regulation. Thyroid illnesses are commonly split into two groups: those that primarily impair thyroid function and those that entail thyroid neoplasms or tumours. In the general population, both sorts of illnesses are rather frequent. The majority of thyroid issues may be properly addressed. Thyroid dysfunction is generally linked with either insufficient hormone production (hypothyroidism) or excessive hormone production (hyperthyroidism). The enlargement of a thyroid nodule in the thyroid gland is a warning indication of thyroid cancer. Thyroid nodules cause no symptoms in the majority of persons. Benign nodules are generally tiny (less than one centimetre in diameter) and require regular monitoring. An ultrasonic (US) imaging technology can detect a thyroid tumour in its early stages. US can assess blood surge to the thyroid and its nodules in addition to volume, position, and number of nodules, distinctness of borders, extra nodule filling such as calcium deposits or the quantity of blood flow. A sample ultrasound image with a thyroid nodule and thyroxine is shown in Figure 1.

Overgrowth of the thyroid gland can result in the formation of one or more nodules. It is unknown why this occurs. When nodules develop, the primary fear is cancer. Fortunately, malignancy is extremely rare, occurring in fewer than 5% of all nodules. People with a family history of nodules and those who do not get enough iodine are more likely to develop nodules. Thyroid hormone requires iodine to be produced. Thyroid nodules can in a variety of shapes and sizes [9]. Colloid nodules: these are thyroid tissue overgrowths that might be one or many in number. These growths are completely safe (not cancer). They can get pretty big, but they never grow bigger than the thyroid gland. Inflammatory nodules: these nodules grow when the thyroid gland is inflamed for a long time (swelling). The pain caused by these growths may or may not be felt. Multinodular goitre: a multinodular goitre is a kind of enlarged thyroid that consists of several nodules (mostly benign). Thyroid nodules that are hyperfunctioning: these nodules produce thyroid hormone on their own, by passing normal feedback control and increasing the risk of hyperthyroidism. Thyroid cancer is rare, with malignant thyroid nodules accounting for less than 5% of all thyroid nodules. This paper focuses on bringing an efficient deep-learning model

for thyroid nodule detection which following are the objectives:

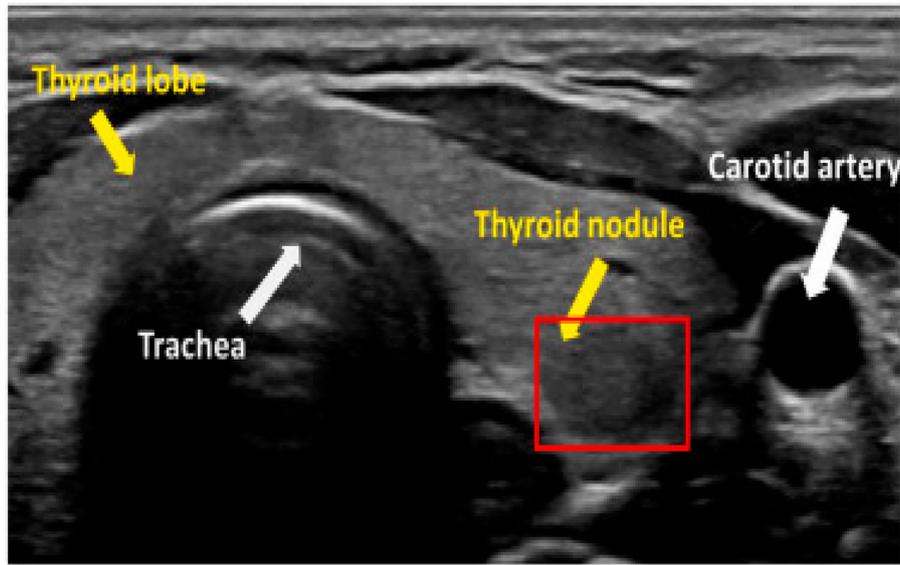
- An efficient deep-learning model is implemented for the classification of thyroid nodules
- Deep learning model is performed over ultrasonic images of the thyroid
- Extraction and selection using GLCM and Multi-objective Particle Swarm with Random Forest Optimization (MbPSRA) for better extraction and dimensionality reduction
- With the help of Residual U-Net architecture, the classification will be done

Paper organization: Section 1 of the study provided an overview of thyroid nodules; the remainder of the paper is as follows. The literature review is presented in Section 2, the methodology is shown in Section 3, the performance analysis is presented in Section 4 and the chapter is concluded in Section 5.

## 2. Literature review

Zhou et al. (2018) [10] used U-Net to differentiate thyroid nodules and presented a segmentation strategy based on it with annotation marks as guidance. To begin, a nodule's four main and minor axis endpoints are manually calculated. Then, at the four places in the picture, four white dots are directly created to aid the deep neural network's training and inference. To differentiate thyroid nodules, Zhou et al. (2018) [10] also used U-Net, and built an interactive segmentation strategy utilizing annotation marks as guidance. A nodule's four main and minor axis endpoints are manually calculated first. Liang et al. (2020) [11] wanted to create a multiorgan CAD system based on CNNs for detecting thyroid and breast issues, as well as see how this system affects the diagnostic effectiveness of various methods of preprocessing. Yang et al. (2021) [12] created a multi-task cascade deep learning model (MCDLM) for automated thyroid nodule detection that uses multimodal ultrasound data and combines radiologists' various topic expertise (DK). To acquire more precise nodule segmentation findings, we transfer U-knowledge networks. The ultrasonic features (UF) of the nodule are then quantified as conditions, which aid in the generation of stronger images and discriminators. Chu et al. (2021) [13] offer a thyroid nodule mark-guided ultrasound deep network segmentation model. The approach utilized in this study achieves similar results when compared to VGG19, Inception V3, DenseNet 161, segmentation accuracy and network operation time.

Edgar Gabriel et al. [5] developed two versions of a code of thyroid FNAC images using texture-based segmentation, which is an important first step towards realizing a fully automated CAD solution.



**Figure 1.** Ultrasound image with thyroid nodule and thyrolobe.

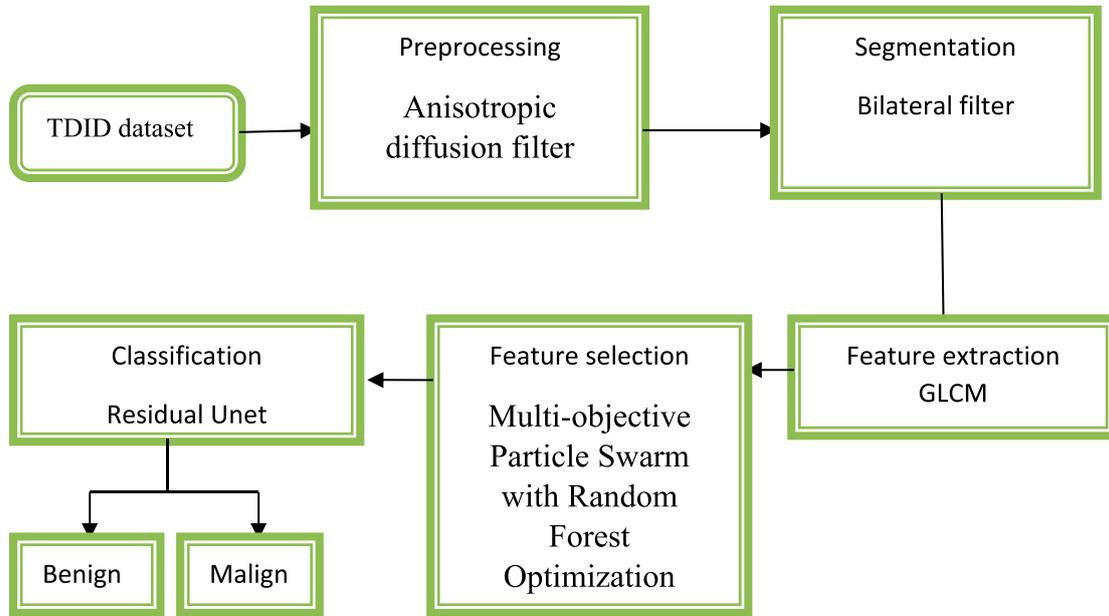
The code has been developed in MPI format to take advantage of distributed memory compute resources. The Variable Background Active Contour algorithm was developed by Maroulis et al. [10] for the detection of abnormal nodules in thyroid images. When compared to the conventional active contour model, the variable background model improved the sensitivity of nodule detection. The authors achieved an average accuracy of 90% for detecting nodules in ultrasound images. To detect and segment abnormal cancer cells in ultrasound thyroid images, Kobayashi et al. [14] used a fuzzy edge detection algorithm. The researchers developed improved generalized fuzzy rules for cancer cell boundary segmentation. Savelonas et al. [6] suggested the variable background active contour model as an active contour model. It is used in ultrasound images to detect thyroid nodules. The new model provides edge independence, less operation smoothing and topological changes. It outperforms the active contour without edges model in terms of accuracy. Accuracy can be increased by using a limited image subset as the background, which changes shape appropriately to decrease the effects of background inhomogeneity.

Preeti Aggarwal et al. [7] introduced a method for automatic segmentation. It contains a summary of all the results obtained using either automatic tools or by applying specific algorithm segmentation on lung CT and thyroid US. For the segmentation of thyroid US images, two tools are available: Analyse 10.0 and Mazda. Tsuda et al. [3] investigated thyroid cancer in Fukushima residents aged 18–26. The ultrasound images were used by the authors to screen for thyroid cancer. While screening thyroid ultrasound images for thyroid cancer, the authors achieved a 95% confidence level. To detect the thyroid nodule in ultrasound images, Nugroho et al. [2] used an active

contour bilateral filtering method. Using this bilateral filtering approach, the irregular boundary of the thyroid nodule was founded and accurately segmented. Before segmentation, the authors found the speckle noises and removed them from the thyroid image. For detecting the thyroid gland in ultrasound images, Gomathy et al. [4] used the principle component analysis method. The thyroid area was accurately segmented using region of interest (ROI) and morphological operations. Du et al. [5] devised a technique for finding thyroid nodules in ultrasound images. An anisotropic diffusion filter (ADF) was used to detect and remove speckle noise. For accurate nodule segmentation, local phase symmetry features were extracted from thyroid images.

### 3. Proposed method

The proposed framework's general design with the steps are listed below. Data extraction from well-known repositories Thyroid Digital Image Database (TDID), which has ultrasound scans of 298 individuals, is then used to gather these images that have been preprocessed from raw files; however, the possibility of noises and anomalies is high. To remove those and enhance the images even better for further stage prediction, the use of ADF has been used. These preprocessed images are passed to the segmentation stage for segmenting the regions using a bilateral filter. Those identified regions will be passed to feature extraction where sufficient features are extracted using GLCM and then from those features, specific features are identified by the feature selection method with the help of MbPSRA. Finally, selected features are passed to the classification stage where with the help of Residual U-Net, it performs and thereby gives the output. Figure 2 displays the overall architecture of the proposed work.



**Figure 2.** Overall architecture of the proposed framework.

**Table 1.** TDID data set description.

Training (benign)	Testing (benign)	Training (malign)	Testing (malign)	Total
41	11	196	50	298

### 3.1. Data collection

The ultrasound image-based thyroid nodule categorization problem has been investigated [15–17], despite the fact that the majority of the data sets utilized in this research are private. Furthermore, due to insufficient time and the particular characteristics of medical diseases, acquiring a large amount of data is exceedingly challenging, necessitating the employment of expensive photo-gathering equipment and patient participation. As a result, we used Pedraza et al. Picture’s Database (TDID), a public thyroid nodule image collection established [15]. In 2015, the TDID data set was made available, which includes ultrasonography thyroid pictures from 298 persons. As a consequence, we were able to collect 450 pictures of thyroid nodules for our research. Each picture assessed by radiologists to determine the state of the thyroid area is given a Thyroid Imaging Reporting And Data System (TI-RADS) score. Thyroid nodule status is assessed using the TI-RADS score, which runs from 1 to 7. It might be any of the numbers below: 1, 2, 3, 4a, 4b, 4c or 5 are the choices. Normal (TI-RADS score 1), benign (TI-RADS score 2) and no abnormal ultrasonography findings are the TI-RADS scores for thyroid nodules with TI-RADS values of 1, 2 and 3. (On the TI-RADS, a score of 3) The letters 4a, 4b, 4c and 5 are used to signify one, two, three or five thyroid nodule characteristics. Thyroid nodule pictures with these four TI-RADS scores have been used to diagnose thyroid nodules in the past (ground-truth labels). The TDID data set is described in Table 1.

### 3.2. Preprocessing

Preprocessing is the most important step in image processing, this stage will clear the noise and anomalies present in the images, enhance the quality, and thereby it will be much more effective in ADF.

#### 3.2.1. Anisotropic diffusion filter

Inhomogeneous regions, such as those around the borders and with few features, AD is adaptable in that it does not employ hard thresholds to change performance [18]. The speckled ultrasound pictures can be improved using AD, however the technique may destroy a few data in the process. Peron and Mallik [19] suggested that nonlinear PDE may be used to smooth an image by implying:

$$\begin{cases} \frac{\partial I}{\partial t} = \text{div}[c(|\nabla I|) \cdot \nabla I] \\ I(t=0) = I_0 \end{cases} \quad (1)$$

where  $\nabla$  is the gradient operator, which recognizes the image’s edge. Furthermore, the authors suggested employing two coefficients:

$$c(x) = \frac{1}{1 + \left(\frac{x}{k}\right)^2} \quad (2)$$

$$c(x) = \exp\left[-\left(\frac{x}{k}\right)^2\right] \quad (3)$$

where the edge magnitude is represented by  $k$

$$|\nabla I| \gg k, \text{ then } c(|\nabla I|) \rightarrow 0 \quad (4)$$

$$|\nabla I| \ll k, \text{ then } c(|\nabla I|) \rightarrow 1 \quad (5)$$

resulting in an Gaussian filtering. A discrete form is denoted by

$$I_s^{t+\Delta t} = I_s^t + \frac{\Delta t}{|\bar{\eta}_s|} \sum_{q \in \bar{\eta}_s} c(\nabla I_{s,p}^t) \nabla I_{s,q}^t \quad (6)$$

where  $I_s^t$  the discretely sampled image, pixel position in a discrete 2D grid denoted by  $s$ , the time step size is denoted  $\Delta t$   $\bar{\eta}_e$  is the spatial neighbourhood of  $s$ ,  $|\bar{\eta}_s|$  is the number of pixels present in the window.

$$\nabla I_{S,q}^t = I_q^t - I_s^t, \forall q \in \bar{\eta}_s \quad (7)$$

Intra-region smoothing and edge maintenance are two of AD's key advantages. For images contaminated by additive noise, AD yields extraordinarily good results [18]. Also, due to its low processing complexity, this approach is favoured [20].

### 3.3. Segmentation

Because as shown in [1,6], lesion form and border are connected to several important parameters for detecting benign and malignant lesions. The borders of most malignant tumours are hazy and uneven. The segmentation process is carried out by a bilateral filter.

#### 3.3.1. Bilateral filter

Bilateral filtering [21] is a low-pass filter that is used to smooth an image by keeping the quality of the object's edge. The effectiveness and reliability of different filters in decreasing speckle are described in [22], whereas the overall equation is as follows:

$$h(q) = \Gamma^{-1}(q) \int a_{\varphi(q)} f(\varepsilon') c(\varepsilon', q) s(f(\varepsilon'), f(q)) d\varepsilon' \quad (8)$$

where

$$\Gamma(q) = \int a_{\varphi(q)} c(\varepsilon', q) s(f(\varepsilon'), f(q)) d\varepsilon' \quad (9)$$

the original image is  $f(q)$ ,  $h(q)$  is the filtered image, a measure of neighbourhood window is denoted by  $Q(q)$ , while  $\varepsilon'$  shows the pixel location.  $(\varepsilon', q)$  and  $s(\varepsilon'), f(q)$ , respectively, defined as

$$c(\varepsilon', q) = \left( \frac{-|q - \varepsilon'|^2}{2\sigma_c^2} \right) \quad (10)$$

$$s(f(\varepsilon'), f(q)) = \exp \left( \frac{(f(q) - f(\varepsilon'))^2}{2\sigma_s^2} \right) \quad (11)$$

where  $\sigma_s$  is the standard deviation of the Gaussian random value and  $\sigma_c$  is the standard deviation for the  $\phi$  window area.

**Table 2.** Illustration of grey levels using GLCM.

Neighbour pixel value	0	1	2	3
0	0,0	0,1	0,2	0,3
1	1,0	1,1	1,2	1,3
2	2,0	2,1	2,2	2,3
3	3,0	3,1	3,2	3,3

### 3.4. Feature extraction

Feature extraction is the process of gathering more detailed information about an image, such as colour, shape and texture. The significant information of a picture is included in features. The types of attributes that features describe are separated into distinct kinds. Texture refers to a distinctive and spatially repeated surface structure created by repeating a single piece or a group of components in different relative spatial places. Texture is a key property for recognizing places of interest in a photograph. One of the first methodologies for extracting texture characteristics was Grey Level Co-occurrence Matrices (GLCM) [23]. After that, the application generates matrices from which statistical measures may be extracted.

The texture of a photograph is its most important feature, and it is frequently employed [24,25]. Texture feature extraction is a crucial stage in the texture analysis process for obtaining this information. Texture attributes may be retrieved in a variety of ways, including structural, statistical, model-based and transform data approaches, with the Gray Level Co-occurrence Matrix being one of the most well known (GLCM). The GLCM maintains second-order statistical information regarding picture pixel spatial connectivity. Haralick derived the Haralick texture features from GLCM by identifying 13 related statistical characteristics. Texture, like clouds and water, is a visual pattern that is not created by the presence of only one colour. Haralick advocated the usage of a co-occurrence matrix with grey levels. It always assesses the connection between two adjacent pixels, the reference pixel being the first and the neighbour pixel being the second. The GLCM matrix is shown in Table 2 for four distinct grey levels. We have utilized a one-pixel offset here (a reference pixel and its immediate neighbour). A bigger offset can be employed when the window is large enough.

### 3.5. Feature selection

A very large number of features are produced by the MbPSRA complexity of an image's content. As we have already mentioned, not all of an image's features are helpful in solving a particular issue, and even those that are can occasionally be redundant. We applied an optimization strategy to only save the most important features in order to solve this issue. The three stages of this strategy, which is based on Multi-Objective Particle Swarm Optimization with Random Forest (MbPSRA),

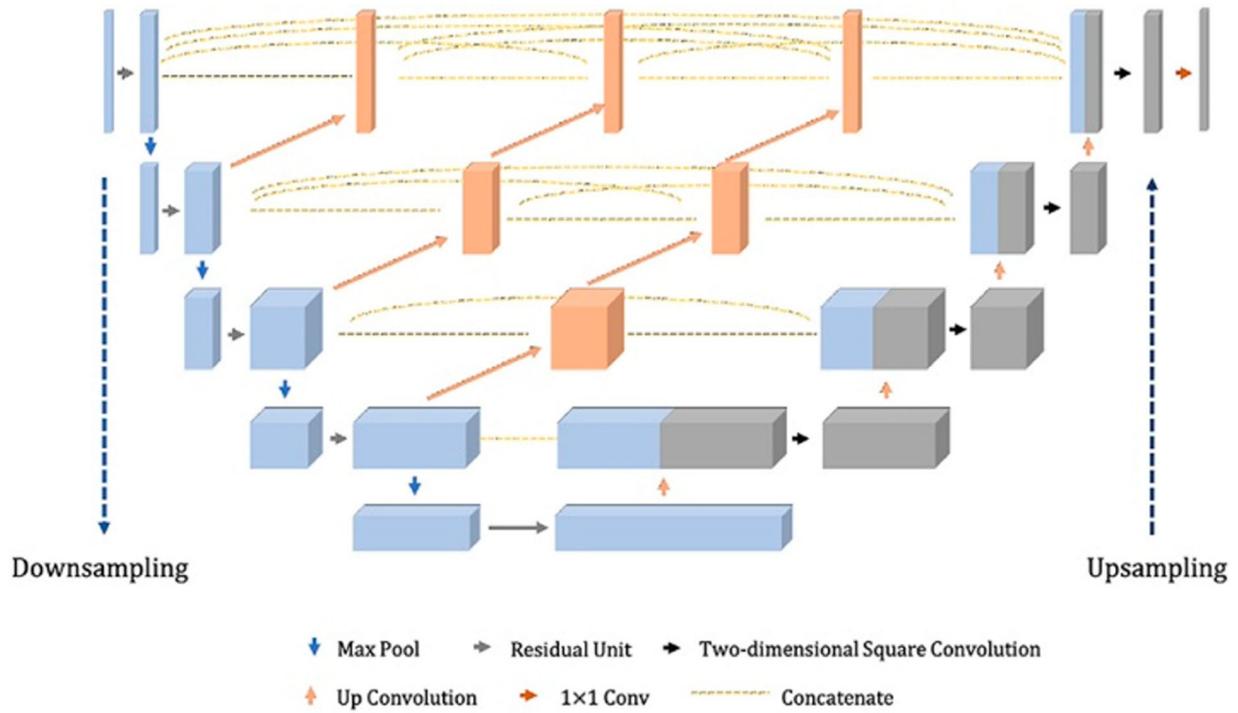


Figure 3. Residual U-Net architecture.

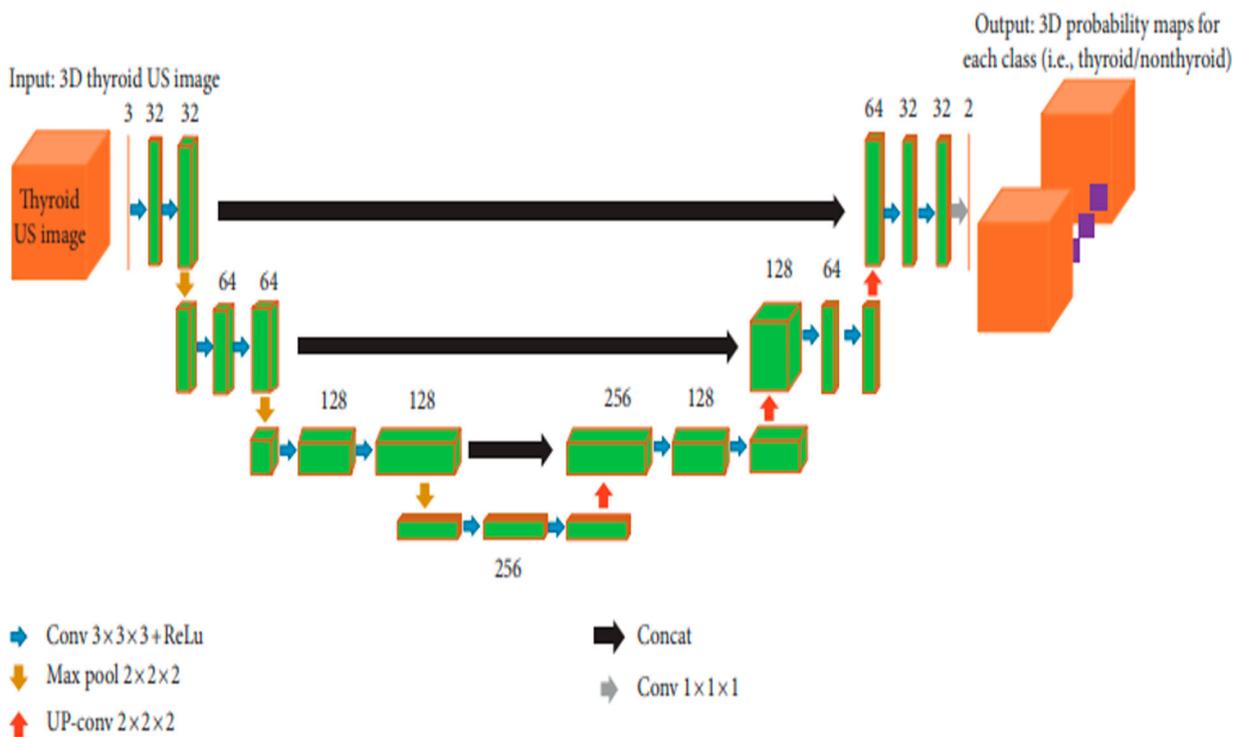


Figure 4. Framework of 3D U-net CNN architecture.

are estimating the significance of the characteristics, assessing redundancy and optimizing.

Measuring feature’s relevance: The effectiveness of a feature in terms of classification is related to the idea of that characteristic’s relevance. We used RFs, which provide an indication of the features that have the greatest effectiveness during a classification, to calculate the relevance value of each feature. A set of binary decision

trees known as an RF is generally more effective than a simple decision tree, but it has the drawback of being more challenging to interpret. A decision tree node for each feature in the RF serves as its representation.

Measuring redundancy: Redundancy and feature correlation are related. The more closely the traits are related, the more redundantly they represent the same information. The dissimilarity matrix of the input

US Feature of Malignant thyroid nodule	Range
Shape– irregular	$\geq 1\text{cm}$
Echo structure solid	$\approx 1\text{cm}$
Internal echo heterogeneous	<b>0.9989</b>

**Figure 5.** Features of a malignant thyroid nodule.

entities is computed using the following correlation coefficient to translate this concept:

$$\sigma(x, y) = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)\text{var}(y)}} \quad (12)$$

$\text{var}()$  stands for a feature's variance, while  $\text{cov}()$  stands for the covariance between two features, where  $x$  and  $y$  are two features. A complete graph is used to model the issue, with the nodes representing the value of each entity's relevance and the edges representing the value of each dissimilarity.

**Optimization:** We build a complete and weighted graph  $G$  from the dissimilarity matrix, where the nodes represent the features and the edges (weight) reflect the similarity between the features. This allows us to optimize the subset of retrieved features. To do this, we use a multi-objective binary PSO  $G$  to locate an ideal sub-graph that only includes the most important attributes. This optimization algorithm has been applied to the extracted features.

### 3.6. Classification

The downsampling path as in Figure 5 on the left successfully extracts image features; the upsampling

**Table 3.** Overall analysis under accuracy, sensitivity and specificity.

Models	Accuracy	Sensitivity	Specificity
CNN	89	92	94
Unet	91	94	96
Attention UNet	86	90	92
RCNN UNet	90	93	95
Nested UNet	92	95	97
ResUNet (Ours)	95	97	98

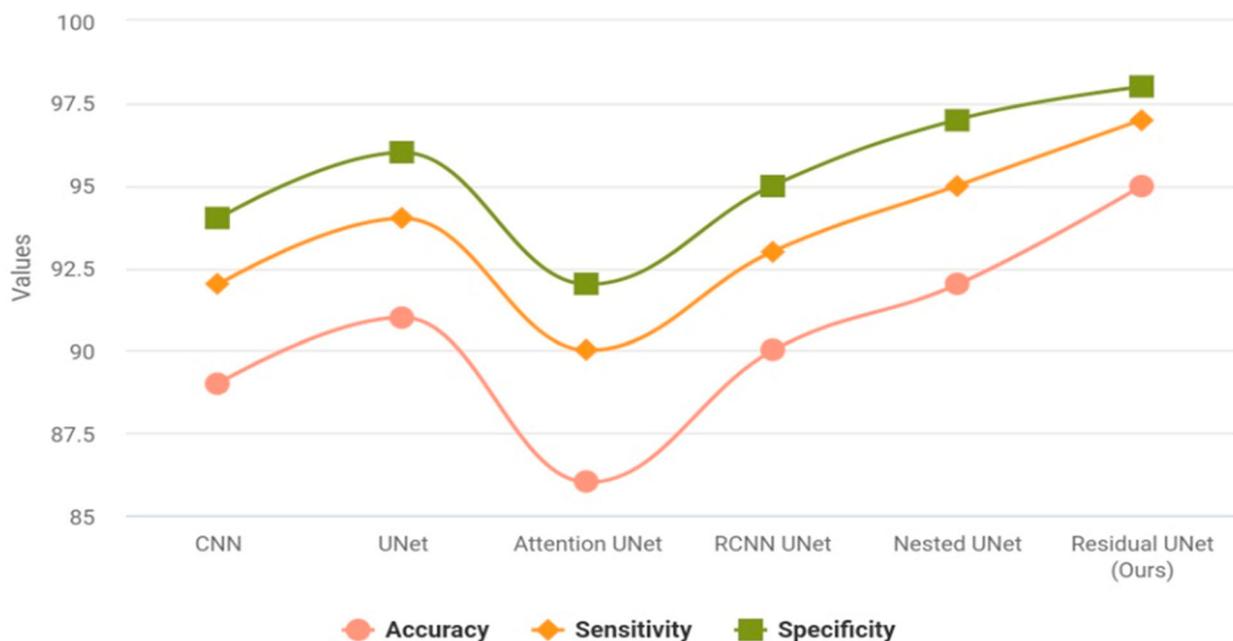
**Table 4.** Overall analysis under precision, recall and F1-score.

Models	Precision	Recall	F1-score
CNN	88	80	87
Unet	89	82	80
AttentionUNet	86	84	86
RCNNUNet	84	81	87
NestedUNet	90	83	89
ResUNet (Ours)	93	85	91

**Table 5.** Overall analysis under detection rate, TPR and FPR.

Models	Detection rate	TPR	FPR
CNN	88	86	14
Unet	89	87	13
AttentionUNet	86	83	17
RCNNUNet	87	84	16
NestedUNet	90	88	12
ResUNet (Ours)	94	95	5

path on the right recovers image size, improves image segmentation accuracy, and allows reconstruction of details through a series of successive transposition convolution operations; and the middle is a series of several convolution operations. Previously, the structure served as a link between the encoder and decoder feature maps, bridging the semantic gap. Figure 3 shows the residual U-Net architecture.



**Figure 6.** Models vs accuracy, sensitivity and specificity.

Although the shallow convolution structure is unable to properly capture the image's complex structure, the stack's deep convolution and redundant structure result in gradient disappearance and tearing. Convolutional blocks must be used in each layer to extract features from the Dense-ResUNet model. To boost model convergence as in Figure 4, a BatchNorm (BN) layer is added to each of these two convolutional units. Feature maps are present in each of the convolutional layers, as shown in Figure 5.

When the convolution kernel scans a pixel in an image and employs detailed content information in the environment to build semantic image features, the activation function (ReLU) characterizes picture characteristics with content and space:

$$X_j^{l+1} = f\left(t_i^{l+1} + \sum_j X_i^l \div k_{ij}^{l+1}\right) \quad (13)$$

where  $X_{j+1}$  shows the feature map,  $X_{il}$  denotes the input feature in the  $(l+1)$ th layer,  $t$  denotes the offset term,  $f$  represents the activation function (rectifier linear unit, ReLU),  $ij$  denotes a collection of input eigenmatrices and  $k$  denotes the convolution kernel. By reducing the size of visual components, the pooling layer can communicate high-level information and semantics:

$$X_j^{l+1} = t_j^{l+1} + X_j^l \otimes k_j^{l+1} \quad (14)$$

where  $\otimes$  refers to the convolutional structure's pooling procedures. Finally, as the prediction layer, the fully connected layer completes the photo categorization assignment by using the maximum likelihood function. Because of its superior learning capabilities, the U-Net model introduced by Ronberger et al. (2015) works

very well in analysing medical pictures containing small samples and difficult modalities.

$$F_{mn} = f\left(s_{mn} + \sum_{i=0}^{a-1} \sum_{j=0}^{b-1} w_{ij} \cdot X_{(m+i)(n+j)}\right) \quad (15)$$

where  $s$  stands for stride,  $w$  for convolution kernel and  $X$  stands for input. The encoder's feature maps pass through a thick convolution block, as can be seen. The number of blocks in the convolution layers is determined by the pyramid level. We suppose that  $X_{ld,lc}$  is a model node, with  $dl$  referring to the encoder's down-sampling layer and  $lc$  referring to the dense block's convolution layer along the skip connection. Meanwhile, we define  $x_{ld,lc}$  as the output of  $X_{ld,lc}$ , then the  $x_{ld,lc}$  can show the feature maps as in Equation (6).

$$x^{l_d, l_c} = \begin{cases} \Delta(x^{l_d-1, l_c}), \\ \Delta([\![x^{l_d, l_c}]_{l_k=0}^{l_c-1} \mathcal{F}(x^{l_d+1, l_c-1})]) \end{cases} \quad (16)$$

where  $\Delta(\cdot)$  stands for the convolution step that follows ReLU,  $M()$  for the upsampling step and  $[\cdot]$  for the concatenation step. It is the pooling layer that merges the data it receives. In this piece we used the maximum pooling approach. Equations (7) and (8) are used to compute the output's height and weight (8).<sup>26,27]</sup>

$$H_{out} = \left\lceil \frac{H_{in} + 2 \times p_i - d_i \times (k_i - 1) - 1}{s_i} + 1 \right\rceil \quad (17)$$

$$W_{out} = \left\lceil \frac{W_{in} + 2 \times p_j - d_j \times (k_j - 1) - 1}{s_j} + 1 \right\rceil \quad (18)$$

The padding is shown by  $p$ , the dilation is represented by  $d$ , the height is shown by  $H$  and the weight

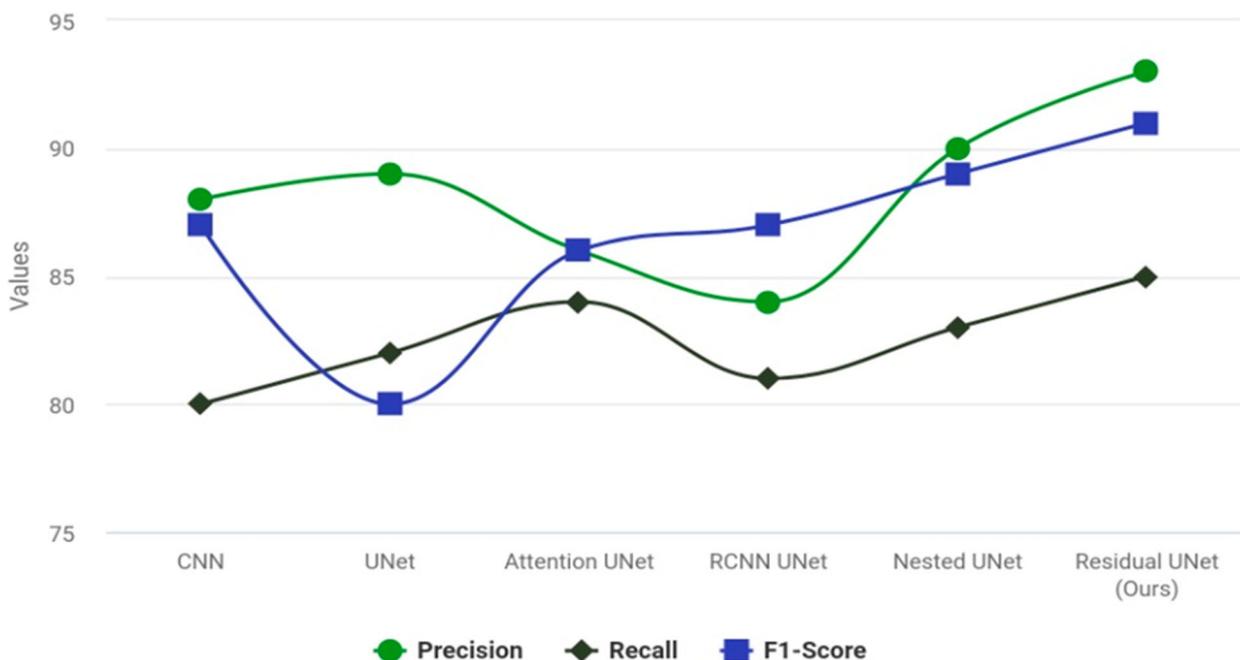
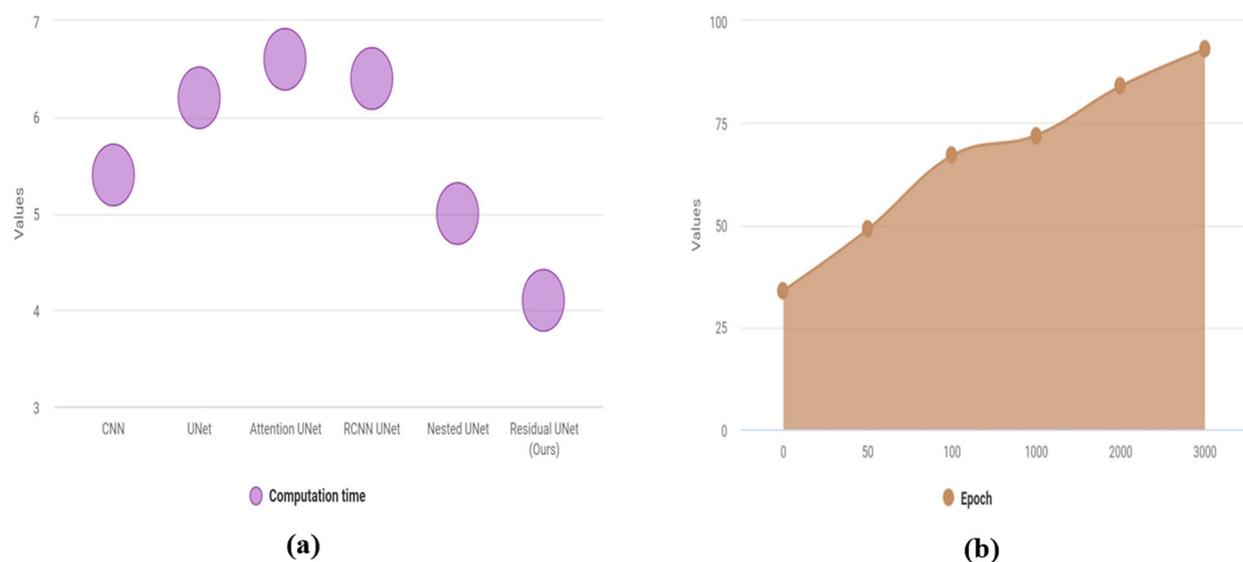


Figure 7. Models vs precision, recall and F1-score.



**Figure 8.** Models vs detection rate, TPR and FPR.



**Figure 9.** (a) Models vs computation time (b) accuracy vs epoch range in percentage wise.

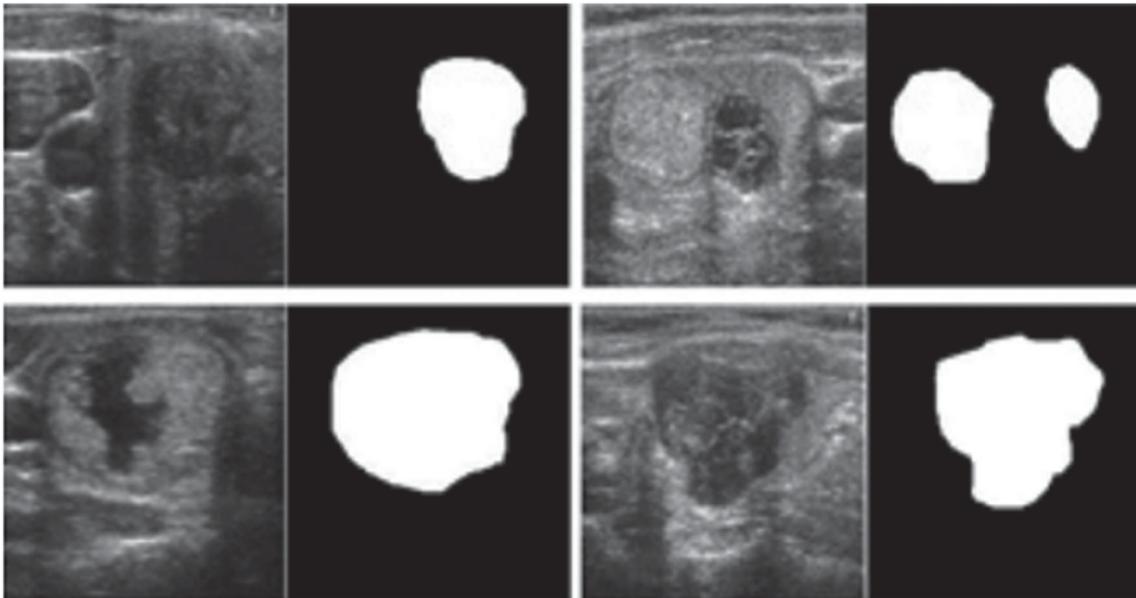
is represented by  $W$ . The network's architecture was created in accordance with the provided requirements. Gaussian noise was applied in the input, coupled with histogram equalization, to make the network contrast independent and linearly scaled to get the Gaussian distribution, similar to UNet. An Adam optimizer was used to train the network, with a learning rate of  $1e-4$  and  $\beta_1 = 0.99$ . As a loss function, binary cross-entropy was applied. With a batch size of 32 and early stopping, the network was trained for 350 epochs.

#### 4. Performance analysis

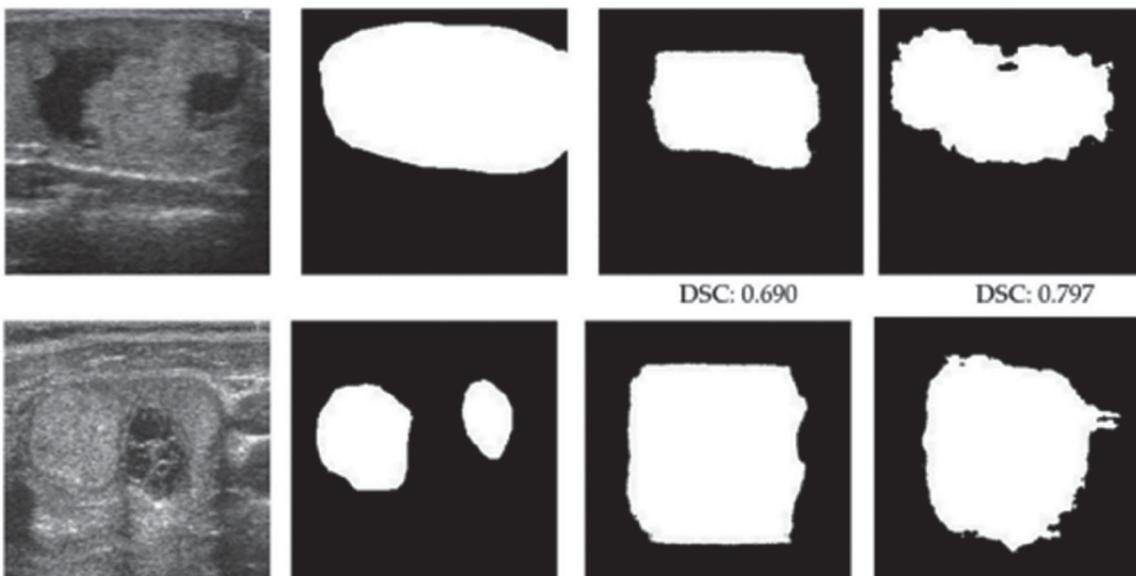
The model is implemented using hardware specifications like Ryzen 5/7 series CPU, 128 GB RAM, 1TB HDD and windows 10 OS, Software specifications like PyTorch an open source python library for developing deep learning models and Google Collaboratory an open source Google environment for building frameworks.

Experiment evaluations are done over various models such as CNN, UNet, Attention UNet, RCNN UNet, Nested UNet over measures like accuracy, sensitivity, specificity, recall, precision, F1-score, detection rate, TPR, FPR and computation time. Tables 3 and 4 depict the overall analysis of models over accuracy, sensitivity and specificity. Figure 6 depicts the graphical representation of various models in which our model outperforms better (accuracy:0.95, sensitivity:0.97, specificity:0.98).

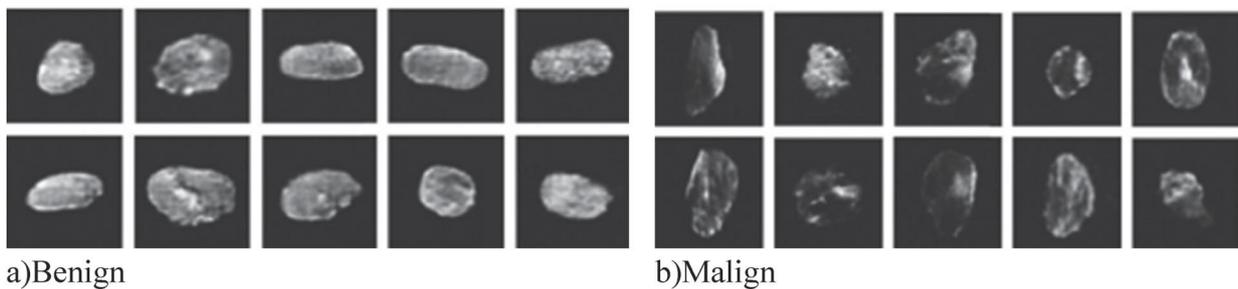
Table 5 depicts the overall analysis under detection rate, TPR and FPR. Figure 7 depicts a graphical representation of various models over the proposed method in which our model outperforms better (detection rate: 0.94, TPR:0.95, FPR:0.5). Figure 8 shows the TPR and FPR, Figure 9(a) shows a graphical representation of different models using the suggested method; our model performs better (FPR:0.5, TPR:0.95, and detection rate: 0.94). Figure 9(b) depicts accuracy vs the



**Figure 10.** Thyroid sample images for the TDID data set.



**Figure 11.** Gives the segmented result of the thyroid lung nodule.



**Figure 12.** Results of synthetic thyroid nodule images for each class. (a) Benign and (b) malign.

range of epochs taken for execution. Figure 10 shows the sample images from the TDID data set.

With a total of 16 3D image volumes, the 3DThyroid dataset was published and designed for thyroid region segmentation.

Our suggested method was used to segment the thyroid region's main region, as seen in the top row of Figure 11. The segmentation accuracy between the TDID and 3DThyroid data sets shows a significant difference. This is because the TDID data set includes a

thyroid that is ill and has a wide range of pixel brightness in that area.

The created images adhere to the accepted terminology because our technology synthesizes the image in the same way as the labels. In other cases, as shown in Figure 12, our method creates synthetic training examples with image textures that are more in line with the texture of the image. On the other hand, radiologists' expertise in their field is quite valuable for identifying diseases. The manual elements that we derive from standardized terminology are currently employed in clinical diagnosis most frequently.

## 5. Conclusion

This paper clearly depicts the importance of thyroid nodule segmentation and classification over ultrasound images. With the help of a deep learning framework, effective detection has taken place. From this paper we clearly gained the procedure of segmenting and classifying the thyroid nodules data's collected from the TDID dataset and segmenting using the active contour model and thereby classifying using the ResUnet model. Experiment results state ResUnet creates a great performance when compared to other models. Also, this paper will be helpful for other research specialists to dig deep learning and generate new integrated models for even more effective results.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## References

- [1] Gaitini D, Evans RM, Ivanac G. Chapter 16: thyroid ultrasound. EFSUMB Course Book. 2011.
- [2] Kwak JY, Han KH, Yoon JH, et al. Thyroid imaging reporting and data system for us features of nodules: a step in establishing better stratification of cancer risk. *Radiology*. 2011;260(3):892–899. doi:10.1148/radiol.11110206
- [3] Lingam RK, Qarib MH, Tolley NS. Evaluating thyroid nodules: predicting and selecting malignant nodules for fine-needle aspiration (fna) cytology. *Insights Imaging*. 2013;4(5):617–624. doi:10.1007/s13244-013-0256-6
- [4] Wang Y, Yue W, Li X, et al. Comparison study of radiomics and deep learning-based methods for thyroid nodules classification using ultrasound images. *Ieee Access*. 2020;8:52010–52017. doi:10.1109/ACCESS.2020.2980290
- [5] Nguyen DT, Pham TD, Batchuluun G, et al. Artificial intelligence-based thyroid nodule classification using information from spatial and frequency domains. *J Clin Med*. 2019;8(11):1976. doi:10.3390/jcm8111976
- [6] Kwon SW, Choi IJ, Kang JY, et al. Ultrasonographic thyroid nodule classification using a deep convolutional neural network with surgical pathology. *J Digit Imaging*. 2020;33(5):1202–1208. doi:10.1007/s10278-020-00362-w
- [7] Gomes Ataide EJ, Ponugoti N, Illanes A, et al. Thyroid nodule classification for physician decision support using machine learning-evaluated geometric and morphological features. *Sensors*. 2020;20(21):6110. doi:10.3390/s20216110
- [8] Baloch ZW, LiVolsi VA. Special types of thyroid carcinoma. *Histopathol*. 2018;72(1):40–52. doi:10.1111/his.13348
- [9] Online source. <https://my.clevelandclinic.org/health/diseases/13121-thyroid-nodule>.
- [10] Zhou S, Wu H, Gong J, et al. Mark-guided segmentation of ultrasonic thyroid nodules using deep learning. *Proceedings of the 2nd International Symposium on Image Computing and Digital Medicine*; 2018, October. p. 21–26.
- [11] Liang X, Yu J, Liao J, et al. Convolutional neural network for breast and thyroid nodules diagnosis in ultrasound imaging. *BioMed Res Int*. 2020;2020.
- [12] Yang W, Dong Y, Du Q, ... Zia MB. Integrate domain knowledge in training multi-task cascade deep learning model for benign–malignant thyroid nodule classification on ultrasound images. *Eng Appl Artif Intell*. 2021;98:104064. doi:10.1016/j.engappai.2020.104064
- [13] Chu C, Zheng J, Zhou Y. Ultrasonic thyroid nodule detection method based on U-Net network. *Comput Methods Programs Biomed*. 2021;199:105906. doi:10.1016/j.cmpb.2020.105906
- [14] Ma J, Wu F, Jiang TA, et al. Ultrasound image-based thyroid nodule automatic segmentation using convolutional neural networks. *Int J Comput Assist Radiol Surg*. 2017;12(11):1895–1910. doi:10.1007/s11548-017-1649-7
- [15] Pedraza L, Vargas C, Narvaez F, et al. An open access thyroid ultrasound-image database. *Proceedings of the 10th International Symposium on Medical Information Processing and Analysis*; 28 January 2015; Cartagena de Indias, Colombia. p. 1–6.
- [16] Zhu Y, Fu Z, Fei J. An image augmentation method using convolutional network for thyroid nodule classification by transfer learning. *Proceedings of the 3rd IEEE International Conference on Computer and Communication*; 13–16 December 2017; Chengdu, China. p. 1819–1823.
- [17] Sudarshan VK, Mookiah MRK, Acharya UR, et al. Application of wavelet techniques for cancer diagnosis using ultrasound images: a review. *Comput Biol Med*. 2016;69:97–111. doi:10.1016/j.compbiomed.2015.12.006
- [18] Yu Y, Acton ST. Speckle reducing anisotropic diffusion. *IEEE Trans Image Process*. November 2002;11(11).
- [19] Perona, Malik J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans Pattern Anal Mach Intell*. 1990;12:629–639. doi:10.1109/34.56205
- [20] Tsiotsios C, Petrou M. On the choice of the parameters for anisotropic diffusion in image processing. *Pattern Recognit*. 2012.
- [21] Tomasi C, Papatrakis CS. (1998). 'Bilateral Filtering for Gray and Color Images,'.
- [22] 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)*; 2015, August: IEEE. p. 43–46.
- [23] Giressha HM, Nanda S. Thyroid nodule segmentation and classification in ultrasound images. *Int J Eng Res Technol*. 2014.

- [24] Song G, Xue F, Zhang C. A model using texture features to differentiate the nature of thyroid nodules on sonography. *J Ultrasound Med.* 2015;34(10):1753–1760. doi:10.7863/ultra.15.14.10045
- [25] Ma J, Wu F, Zhu J, et al. A pre-trained convolutional neural network based method for thyroid nodule diagnosis. *Ultrasonics.* 2017;73:221–230. doi:10.1016/j.ultras.2016.09.011
- [26] Pan H, Zhou Q, Latecki LJ. SGUNET: semantic guided UNET for thyroid nodule segmentation. 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI); 2021, April: IEEE. p. 630–634.
- [27] Shahroudnejad A, Vega R, Forouzandeh A, ... Punithakumar K. Thyroid nodule segmentation and classification using deep convolutional neural network and rule-based classifiers. 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); 2021, November: IEEE. p. 3118–3121.