






Swin Transformer and Attention Guided Thyroid Nodule Segmentation on Ultrasound Images

Arunkumar Beyyala^{1*}, Rajaram Priya¹, Subramani Roy Choudari², Rajaram Bhavani¹

¹ Department of Computer Science and Engineering, Annamalai University, Chidambaram 608002, Tamilnadu, India

² Department of Computer Science and Engineering, Usha Rama College of Engineering and Technology, Gannavaram 521109, Andhra Pradesh, India

Corresponding Author Email: arunbeyyala@gmail.com

Copyright: ©2024 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.290109>

ABSTRACT

Received: 25 September 2023

Revised: 16 January 2024

Accepted: 23 January 2024

Available online: 27 February 2024

Keywords:

attention guided network, DDTI, deep learning, Swin Transformer, thyroid nodule, ultrasound image segmentation

The early detection of thyroid cancers depends on the ability to segment thyroid nodules. Thyroid ultrasound imaging is critical for diagnosing thyroid nodules. It is very challenging to correctly segment and extract the thyroid nodule from ultrasound pictures. To solve this issue, a deep learning-based segmentation module called Thyroid Region Prior Guided is recommended for the first segmentation process. This module uses three different forms of encoder, decoder, and thyroid region previous guidance to prepare for the nodule picture. This method uses ultrasound to precisely segment thyroid nodules utilizing the Swin Transformer and Attention Guided network. The DDTI dataset is used to assess the effectiveness and performance of the proposed model. The performance of the proposed model is evaluated using a number of metrics, such as accuracy of 83.43, F1-score of 68.35, IoU of 59.67, and Dice of 71.87, which ensures the best accuracy improvement over the models currently in use.

1. INTRODUCTION

The most widely used technologies today are medical imaging and artificial intelligence (AI). AI is most frequently used in research facilities to process medical image data. Ultrasonic image segmentation is the technique that is used in research the most frequently [1]. One of the most significant endocrine organs in the human body is the thyroid gland. This has the power to control metabolism, growth, and thyroid hormone production. Thyroid nodules therefore frequently cause endocrine system disorders in the adult population [2]. In the preceding five years, 53,000 new cases of thyroid cancer were identified and diagnosed just in the United States. Thyroid cancer is one of the cancers that progresses the quickest of all. The following key contributions are included in the proposed work:

- The multi-task deep learning-based attention guided network is proposed to thyroid nodule and gland regions segmentations, and then process to identify the nodule location within the gland region.
- In the encoder, avoiding the overfitting by using Size Prediction Task (SPT) and set up the fusion feature between the nodule and gland prior future map.
- The Swin Transformer network is included in the prediction of the attention guided for the supervision learning, which is controllable and interpretable.
- The fusion feature is used to improve the channel effectiveness and achieving the better performance of the attention model.

2. LITERATURE SURVEY

A model for thyroid nodule image segmentation in medical imaging was described by Liu et al. [3] and was built on an optimised U-Net Convolutional Neural Network. In order to structure the segmentation process, the Test Time Augmentation (TTA) approach was applied to the U-Net. The segmentation effect was improved by the segmentation procedure using the TTA approach, which did not require a cascade network. Although the model was time-intensive, the current method obtains the highest levels of accuracy and picture segmentation results while also identifying ultrasound in low-quality images.

The thyroid nodule is one of the most crucial ultrasound images, and it is quite useful for early diagnosis of people with the ailment. This has proven feasible in previous decades at a pace of 4.5% annually. In the previous five years, 53,000 new cases of thyroid cancer were found and diagnosed only in America. Nearly 2000 people died as a result of thyroid cancer [3]. It is more challenging to distinguish the nodules from the normal tissues when they are clearly apparent. In order to use thyroid ultrasonography in Computer-Aided Diagnosis (CAD), an automatic segmentation approach must be developed [4]. Nodules are difficult to recognise automatically because of their unpredictable shapes, jagged edges, and hazy appearances [5].

Ultrasonography makes the thyroid lesions visible, which enables the radiologist to correctly diagnose any conditions. Big thyroid nodules can mutate and turn cancerous in addition to impacting other organs like the trachea and blood vessels

[6]. It is difficult to separate nodules from the images when there is low contrast, signal loss, shadowing, and speckles in the ultrasound images [7]. The key and basic stage in CAD systems to automatically segment the nodular image is the nodule segmentation. For the development of a trustworthy diagnosis system and for ultrasonography nodule biopsies, the automatic segmentation of thyroid nodules is essential [8]. Clinicians can determine a nodule's malignant or benign status depending on its shape [9]. It is very common to segment thyroid pictures using a Convolutional Neural Network-based network model. CNN characteristics are limited by two properties of convolutional operations: inner locality and dependence on the immediate environment [10].

Liu et al. [11] proposed a model for thyroid nodule image segmentation in medical imaging that was based on an enhanced U-Net convolutional neural network. Gong et al. [12] presented a multi-task learning architecture of the thyroid region prior guided feature enhancement network (TRFE-Net) to segregate the thyroid nodule using ultrasound images. Using the multiple RPG modules, this method expanded the nodule segmentation's features. Pan et al. [13] presented a Semantic Guided UNet (SGUNet) for segmenting the thyroid nodule in accordance with the network design. This model used the semantic map of a single channel pixel and presented the low semantic characteristics from the high-level in each stage of decoding in order to access the best nodule accuracy.

Buda et al. [14] reported a segmentation utilising deep learning in two separate methods to locate the callipers in thyroid nodules of image processing. The model that was given had the advantage of producing the best performance and accuracy results, but the procedure necessitates training on a large number of thyroid nodule-specific images.

A deep learning-based ultrasound mark-guided network model for the detection of thyroid nodules was described by Chu et al. [15] using the U-Net network. Superior segmentation accuracy, more application possibilities, and reduced picture needs for model training are all benefits of this approach. For the limits of that model, the single channel and

low contrast were not optimal.

The related efforts have some disadvantages, such as low contrast, slow convergence speed, needing a large number of images for testing and training, temporal complexity, the inability to predict a single channel, and costly computational networks. The segmentation model produces inaccurate results as a result of these drawbacks [16]. A Deep Learning-based attention-directed model is offered as a solution for the segmentation of thyroid nodules.

3. PROPOSED METHOD

The segmentation of human thyroid nodules is improved in this work using Group Normalization-based Deep Learning-based previous directed thyroid region feature segmentation. Before teaching picture segmentation, this framework combines the three stages of encoder, decoder, and thyroid area module. Swin-Transformer's Attention Guided is trained to foresee the thyroid nodule using these sorts. We choose one image from each dataset to create a mini-batch during the training process; the nodule image and gland image are indicated in blue block and green block, respectively.

In Figure 1, the nodule images are delivered to the inferior branch for loss back-propagation while the gland images are supplied into the top branch. We explore the interaction between the gland region prior and the nodule segmentation job in further detail to make better use of the thyroid region prior. We learn that not all nodules require the thyroid gland's care. The thyroid glands' prior knowledge may be damaging in this case. Therefore, in order to make the encoder aware of the nodule size, we suggested a Size Prediction Task (SPT). In particular, we add a size prediction module to the encoder after the fifth layer. To avoid the over-fitting of the model, we adopt a smooth version of the MSE loss (i.e., Logcosh loss) Natekin and Knoll [17] to calculate the loss of the Size Prediction Task Lsize.

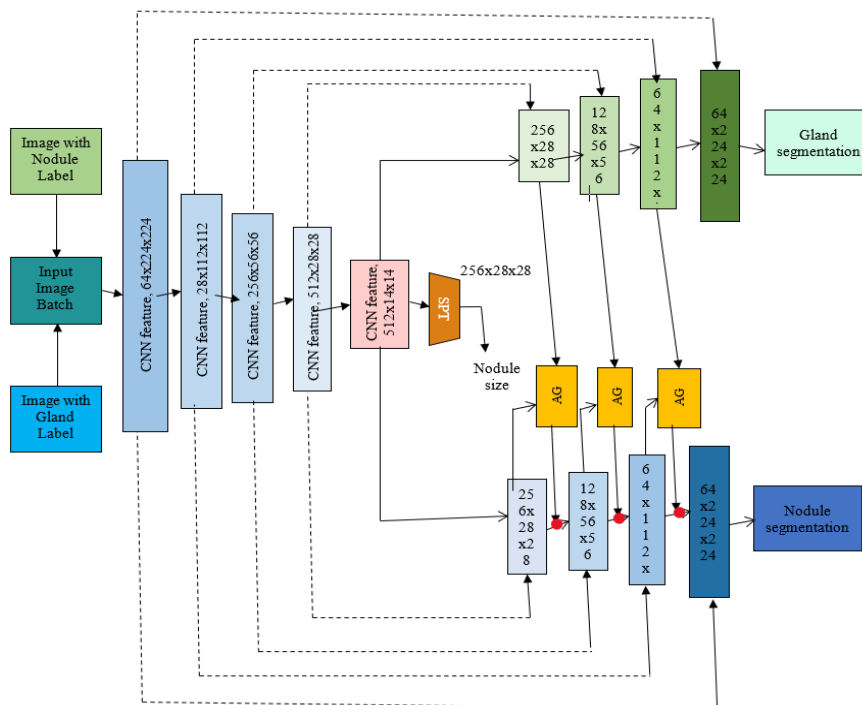


Figure 1. Proposed method of thyroid region prior attention guided segmentation

It is suggested how to segment the thyroid nodule in ultrasound pictures utilising the previously guided feature of the thyroid region-based Attention Guided network model. Prior guidance for the thyroid area, encoders, and decoders are the three main categories of the network model. Images of thyroid nodules are utilised to extract the high-dimensional image feature using the encoder. An essential step in filtering the features is the decoder task. thyroid nodule recognition, thyroid nodule segmentation, and nodule size forecasting. The capability of the preceding guiding module is used to improve the effectiveness of segmenting nodules in past knowledge.

The previously guided feature of the thyroid area is advised to be used with the Digital Database Thyroid Image (DDTI) dataset. The DDTI is a user-friendly, compact, and easily accessible tool for ultrasound department users. The DDTI has all 637 ultrasound images from around 300 patients. On a single device, these images have collected a label of pixel-by-pixel lesions mask. This dataset can be used as an external test set to evaluate deep learning model-based algorithms' performance and room for improvement. The thyroid nodule can be correctly diagnosed with this dataset. This dataset can be used to train and test the segmentation of thyroid glands and nodules. In order to properly segment the image, the Convolutional Neural Network (CNN) approach is utilized for pixel-by-pixel feature extraction and image identification of the thyroid images

3.1 Multi-task deep learning-based attention guided network

A restricted selection of RREMI images serves as the method's sole input image. Figure 2 divides the pictures of the thyroid gland and nodule into green and blue squares. Images of thyroid nodule and gland are segmented using a large number of deep neural networks. The gland and nodule input, which can be of any size, can be used by the CNN to predict the spatial density. In the encoder and decoder portions of the block, it was produced utilising convolutional filters, pooling, batch normalisation, and deconvolutional layers. The convolutional and deconvolutional layers contain a range of filters, including 28, 64, 128, 256, and 512. The concatenation procedure is used to connect the blocks of the encoder and decoder. The attention-guided module receives the features from these photos to increase the nodule segmentation accuracy. The thyroid module's network-based segmentation was done using the UNet encoder and decoder [17]. The refinement layer is used to finish the nodule segmentation procedure that the decoder began.

3.2 Size Prediction Task (SPT) - Encoder

The interaction between the thyroid nodule and the gland segmentation process is thoroughly investigated in order to improve the prior guiding module. If a gland block is found, searching for larger nodules may be of concern because the size of the nodule and the gland are incompatible. These events have had a detrimental effect on prior understanding of the thyroid gland [18]. The Size Prediction Task (SPT), which informs the encoder of the nodule's size, solves this issue. The fifth layer encoder can be extensively included in the SPT. Multiple layer Perceptron (MLP) projected size can be gathered by:

$$S^{pred} = \text{sigmoid}(MLP(f5)) \quad (1)$$

where, S^{pred} is the size prediction. sigmoid is the function of sigmoid. $f5$ is the fifth layer in feature map.

The size prediction loss is calculated to avoid the model of overfitting, this method uses the MSE loss smooth version.

$$L_{size} = \frac{1}{N_{nodule}} \sum_{i=1}^{N_{nodule}} \log(\cosh(S_i^{pred} - S_i^{gt})) \quad (2)$$

where, L_{size} - Size Prediction Task N_{nodule} - number of nodule images S^{gt} - ratio of area of ground-truth and nodule engaged in the image i - decoder feature map index \log - log function and \cosh - hyperbolic cosine function.

3.3 Swin Transformer and attention guided decoder

Combining the Swin Transformer and the Attention Guided decoder significantly improves the multi-task learning framework. The Transformer uses shifted windows to implement the attention guided components. Because the network of prior attention has limited capability in the multi-task feature and retrieving the information of border, Swin Transformer can be added to the network to increase modelling power. Layer Norm layers (LN), multilayer perceptron (MLP), and multi-head self-attention (MSA) modules are all continuously connected two times for each module. The first module of the Swin Transformer uses the window-based MSA (W-MSA) module, and the second module uses the shifted-window-based MSA (SW-MSA). The process of learning attention of the Swin Transformer with the conventional and shifted window can be formulated as:

$$\hat{z}^l = W - MSA(LN(Z^{l-1})) + z^{l-1} \quad (3)$$

$$\hat{z}^{l+1} = SW - MSA(LN(Z^l)) + z^l \quad (4)$$

$$\text{Attention} = \text{softmax}\left(\frac{QK^T}{\sqrt{d}} + B\right)V \quad (5)$$

The ROI (Thyroid) pictures are identified during the segmentation process, which is the first stage of this cascaded network. By using this method, the attention guidance decoder can execute the prior attention network. This method employs an attention-guided decoder to anticipate ROI-related attention using the output of the Swin Transformer. In the RPG module by taking both the channel relationship and the position information into account. The new tailor-designed Adaptive Region Prior Guidance (ARPG) module is shown in Figure 2 (Attention Guiding Decoder-The illustration of the attention guiding decoder. It receives the multi-scale features from Swin Transformer blocks, and then performs attention feature fusing and up-sampling, which plays a role of refining the feature representations and improving the quantity of segmentation.).

3.4 Fusion feature

The decoder job, which can be used to concentrate attention on particular blocks to improve efficiency and give clearer segmentation, is an essential component of filtering the features. The feature map between E_1 and E_4 for the fusion feature is extracted using the block of the one-to-one transformer and then sent to the decoder of the directing attention.

The nodules in the nodule images vary in size and have ambiguous spatial locations. The earlier techniques often segment or crop the thyroid nodule regions at first, allowing

them a chance to benefit from context-dependent error propagation and information loss. Since the encoded feature map has varying spatial size, the upsampling capability of the bilinear interpolation is initially utilised to clarify the equal spatial size. The cost of computing is lowered after the feature map is limited to irrelevant data [19]. The encoder feature maps were ultimately combined in order to generate the prediction of attention after the convolutional layer with the sigmoid.

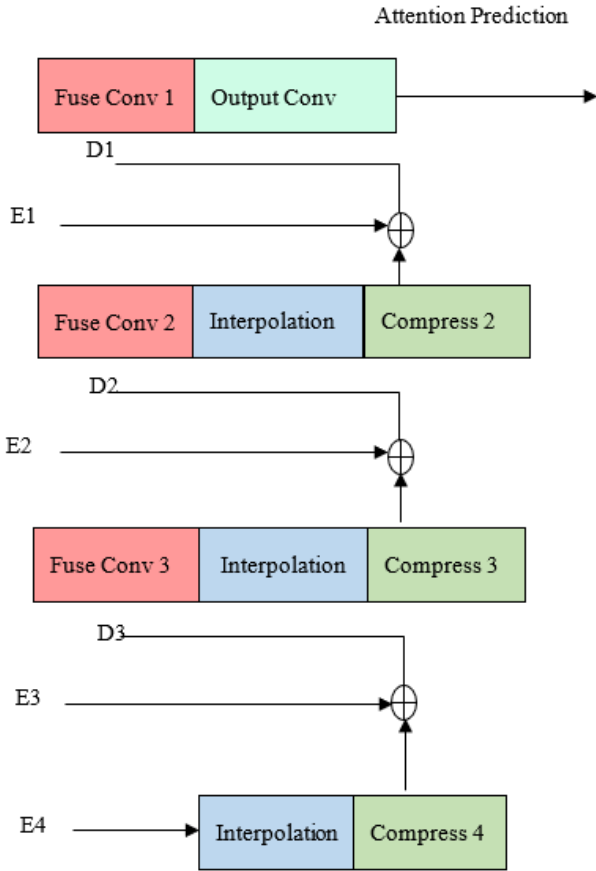


Figure 2. Attention Guiding Decoder - Fusion and up-sampling of multi-scale features for refined segmentation

In this approach, the shallowest and deepest features of the thyroid RREMI image, E_1 and E_4 , are used to describe the retrieved feature maps from the Swin Transformer output [20]. The earlier techniques often segment or crop the thyroid nodule regions at first, allowing them a chance to benefit from context-dependent error propagation and information loss. The decoder feature can be calculated as,

$$D_3 = W_{c4}^T E_4 \oplus E_3 \quad (6)$$

$$D_2 = (W_{c3}^T (W_3^T D_3)) \oplus E_2 \quad (7)$$

$$D_1 = (W_{c2}^T (W_2^T D_2)) \oplus E_1 \quad (8)$$

where, D_3 describes the combination of features of E_4 and E_3 fused future module D_2 describes the combination of features of E_3 and E_2 fused future module D_1 describes the combination of features of E_2 and E_1 fused future module W_{c4} , W_{c3} and W_{c2} describes the corresponding squeezed convolutions used for segmentation and depth estimation in the dimension of the channel. W_4 , W_3 and W_2 are the fusion

convolutions for integration of the information combined within images and \oplus is the concatenation feature operation.

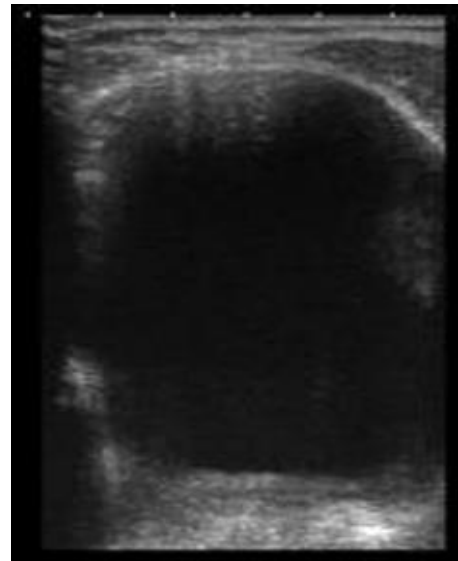
At guiding decoder ending, the output Y is calculated.

$$Y = \sigma(W_{out}^T (W_1^T D_1)) \quad (9)$$

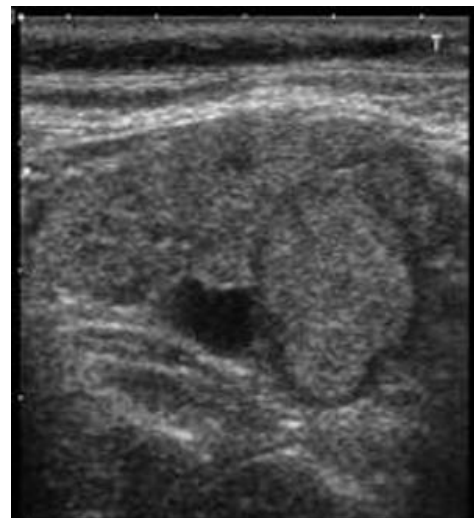
where, W_1 is the fuse convolution to fuse E_2 and E_1 , W_{out}^T is the convolution of the output, and σ is the sigmoid activation operation.

3.5 Dataset

The DDTI dataset, which is provided by Columbia National University, is used in this study to classify thyroid nodules. This collection includes 480 US pictures from 400 thyroid illness patients. After ignoring the 480 faulty images, 427 US images are all that are left. Based on the TI-RADS standard, the thyroid pictures in this dataset are divided into six classes: 2, 3, 4A, 4B, 4C, and 5. The output images of the RREMI [21, 22] algorithm and photos of normal glands are taken into consideration in that. Figures 3.a and 3.b, respectively, depict sample pictures of nodules.



(a)



(b)

Figure 3. Sample images, a) Output of RREMI images, b) Gland Images

3.6 Group normalization

The channels are divided into groups using the group normalisation procedure, which then normalises the features of each group. It classifies pre-defined parameters into groups and splits the channel generally. The setting of multi-task learning has revealed an inherent issue with the RREMI images. Images subsequently shrink in size and appear in a small block; though this may be obvious, the domain change should make the default Batch Norm useless. Therefore, it is advised to swap out the Batch Norm operation for the simple, effective Group Norm operation, which can be used to reduce the difficulty of optimisation and make deep neural network convergence easier. However, the neural networks may not perform as well if a batch's samples come from a different domain or if there are few samples in a batch. The little block's samples with various domain affects can be removed because this Group Norm operation will normalise the block in contrast to the channel C perspective. Faster convergence rate and less framework loss were produced by the Group Norm's operation.

4. EXPERIMENTAL RESULTS

However, the performance of the neural networks may be compromised if a batch's samples come from a different domain or if there are few examples in a batch. By normalising the block from the perspective of channel C, this Group Norm operation can exclude the samples in the small block that are influenced by various domains. As a result of the Group Norm's operation, the convergence rate was quicker and the framework loss was lower.:

- **Accuracy:** The ratio of all correct classifications to the total number of the classifications.

$$Accuracy = \frac{TN+TP}{TN+TP+FN+FP}$$

- **Precision:** The ratio of true positive over the classifications of all positive.

$$Precision = \frac{TP}{TP+FP}$$

- **F1-score:** Combines precision and recall gives average value of weight.

$$F1 - score = \frac{2 * PR * RE}{PR + RE}$$

- **Intersection Over Union:** Extent of overlap of two boxes.

$$IoU = \frac{TP}{FP+FN}$$

- **Dice Coefficient:** Measures the similarities between two sets of data.

$$DICE = \frac{2 * TP}{FP+FN+2 * TP}$$

where, TP, FP, TN, FN, PR, RE indicate true positive, false positive, true negative, and false negative, precision, recall

respectively.

4.1 Performance measures

This section shows the performance metrics for the Convolutional Neural Network model in terms of the potential sum rate. The CNN model was used to run the experimental U-Net, SGUNet, and TRFE+ models. The experimental results of various deep learning models are shown in Table 1 using evaluation metrics like accuracy, F1-score, IoU, and Dice.

Table 1. Represents the performance of various network models

Methods	Accuracy (%)	F1-score (%)	IoU (%)	Dice (%)
U-Net	30.90	49.53	38.98	50.56
SGUNet	80.80	51.34	41.86	53.45
TRFE+	81.81	62.65	51.70	66.74
Attention Guided	83.43	68.35	59.67	71.87

The performance metrics for several models are shown in Figure 4. On the basis of accuracy, F1-score, Intersection Over Union, and Dice coefficient, the proposed Attention Guided Network model is evaluated and compared with the U-Net, SGUNet, and TRFE+ network techniques.

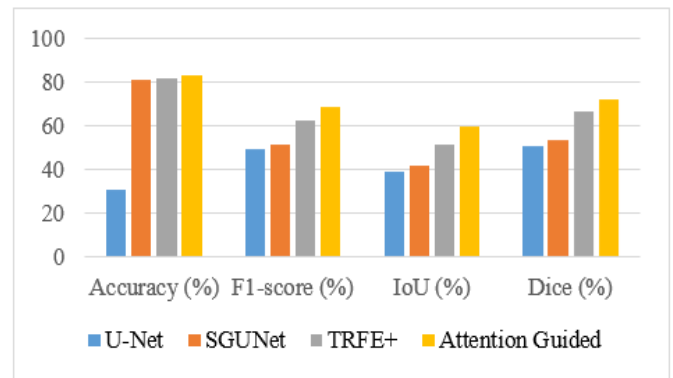


Figure 4. Represents the performance of the various network models

5. CONCLUSION

In this study, the preceding thyroid nodule was segmented using the network-named attention-guided framework. This strategy is built as a multi-task learning algorithm of Convolutional Neural Network, which can train to segment the gland and nodule region and forecast the size of the nodule by using a technique like SPT. During the segmentation process, these techniques enable accurate nodule and gland picture prediction and segmentation. To improve precise segmentation, the previously proposed thyroid region module makes use of the Attention Guided and Swin Transformer decoder modules. In the experimental results, the DDTI dataset provides more precise segmentation and prediction. The suggested model successfully completes the classification task with an accuracy score of 83.43, F1-score of 68.35, IoU of 59.67, and Dice of 71.87. Using the feature selection process in future work can improve the accuracy of the prior thyroid nodule segmentation model. These models exhibited

better diagnostic performance than state-of-the-art models. Based on the performance of different convolutional neural network models, the proposed approach can significantly improve the diagnosing capability of CAD systems for thyroid nodules. Furthermore, the model represents a generalized platform that can assist clinicians working across multiple domains.

REFERENCES

- [1] Wu, Y., Shen, X., Bu, F., Tian, J. (2020). Ultrasound image segmentation method for thyroid nodules using ASPP fusion features. *IEEE Access*, 8: 172457-172466. <https://doi.org/10.1109/ACCESS.2020.3022249>
- [2] Tao, Z., Dang, H., Shi, Y., Wang, W., Wang, X., Ren, S. (2022). Local and context-attention adaptive LCA-net for thyroid nodule segmentation in ultrasound images. *Sensors*, 22(16): 5984. <https://doi.org/10.3390/s22165984>
- [3] Liu, M., Yuan, X., Zhang, Y., Chang, K., Deng, Z., Xue, J. (2020). An end to end thyroid nodule segmentation model based on optimized U-net convolutional neural network. In *Proceedings of the 1st International Symposium on Artificial Intelligence in Medical Sciences*, pp. 74-78. <https://doi.org/10.1145/3429889.3429903>
- [4] Kumar, V., Webb, J., Gregory, A., Meixner, D.D., Knudsen, J.M., Callstrom, M., Fatemi, M., Alizad, A. (2020). Automated segmentation of thyroid nodule, gland, and cystic components from ultrasound images using deep learning. *IEEE Access*, 8: 63482-63496. <https://doi.org/10.1109/access.2020.2982390>
- [5] Koundal, D., Sharma, B., Guo, Y. (2020). Intuitionistic based segmentation of thyroid nodules in ultrasound images. *Computers in Biology and Medicine*, 121: 103776. <https://doi.org/10.1016/j.compbiomed.2020.103776>
- [6] Abdolali, F., Kapur, J., Jaremko, J.L., Noga, M., Hareendranathan, A.R., Punithakumar, K. (2020). Automated thyroid nodule detection from ultrasound imaging using deep convolutional neural networks. *Computers in Biology and Medicine*, 122: 103871. <https://doi.org/10.1016/j.compbiomed.2020.103871>
- [7] Kunapinun, A., Dailey, M.N., Songsaeng, D., Parnichkun, M., Keatmanee, C., Ekpanyapong, M. (2023). Improving GAN learning dynamics for thyroid nodule segmentation. *Ultrasound in Medicine & Biology*, 49(2): 416-430. <https://doi.org/10.3390/s23167289>
- [8] Nguyen, D.T., Choi, J., Park, K.R. (2022). Thyroid nodule segmentation in ultrasound image based on information fusion of suggestion and enhancement networks. *Mathematics*, 10(19): 3484. <https://doi.org/10.3390/math10193484>
- [9] Abbasian Ardakani, A., Bitarafan-Rajabi, A., Mohammadzadeh, A., Mohammadi, A., Riazi, R., Abolghasemi, J., Homayoun Jafari, A., Bagher Shiran, M. (2019). A hybrid multilayer filtering approach for thyroid nodule segmentation on ultrasound images. *Journal of Ultrasound in Medicine*, 38(3): 629-640. <https://doi.org/10.1002/jum.14731>
- [10] Gong, H., Chen, J., Chen, G., Li, H., Li, G., Chen, F. (2023). Thyroid region prior guided attention for ultrasound segmentation of thyroid nodules. *Computers in Biology and Medicine*, 155: 106389. <https://doi.org/10.1016/j.compbiomed.2022.106389>
- [11] Liao, Z., Xu, K., Fan, N. (2022). Swin Transformer assisted prior attention network for medical image segmentation. In *Proceedings of the 8th International Conference on Computing and Artificial Intelligence*, pp. 491-497. <https://doi.org/10.1145/3532213.3532287>
- [12] Gong, H., Chen, G., Wang, R., Xie, X., Mao, M., Yu, Y., Chen, F., Li, G. (2021). Multi-task learning for thyroid nodule segmentation with thyroid region prior. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pp. 257-261. <https://doi.org/10.1109/ISBI48211.2021.9434087>
- [13] Pan, H., Zhou, Q., Latecki, L.J. (2021). Sgunet: Semantic guided unet for thyroid nodule segmentation. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pp. 630-634. <https://doi.org/10.1109/ISBI48211.2021.9434051>
- [14] Buda, M., Wildman-Tobriner, B., Castor, K., Hoang, J.K., Mazurowski, M.A. (2020). Deep learning-based segmentation of nodules in thyroid ultrasound: improving performance by utilizing markers present in the images. *Ultrasound in Medicine & Biology*, 46(2): 415-421. <https://doi.org/10.1016/j.ultrasmedbio.2019.10.003>
- [15] Chu, C., Zheng, J., Zhou, Y. (2021). Ultrasonic thyroid nodule detection method based on U-Net network. *Computer Methods and Programs in Biomedicine*, 199: 105906. <https://doi.org/10.1016/j.cmpb.2020.105906>
- [16] Wang, B., Yuan, F., Lv, Z., He, Y., Chen, Z., Hu, J., Zheng, S., Liu, H. (2022). Hierarchical deep learning networks for classification of ultrasonic thyroid nodules. *Journal of Imaging Science and Technology*, 66(4): 040409-1-040409-10.
- [17] Natekin, A., Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*, 7: 21. <https://doi.org/10.3389/fnbot.2013.00021>
- [18] Manh, V.T., Zhou, J., Jia, X., Lin, Z., Xu, W., Mei, Z., Dong, Y., Yang, X., Huang, R., Ni, D. (2022). Multi-attribute attention network for interpretable diagnosis of thyroid nodules in ultrasound images. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 69(9): 2611-2620. <https://doi.org/10.1109/TUFFC.2022.3190012>
- [19] Yang, W., Dong, Y., Du, Q., Qiang, Y., Wu, K., Zhao, J., Yang, X., Zia, M.B. (2021). Integrate domain knowledge in training multi-task cascade deep learning model for benign-malignant thyroid nodule classification on ultrasound images. *Engineering Applications of Artificial Intelligence*, 98: 104064. <https://doi.org/10.1016/j.engappai.2020.104064>
- [20] Liu, T., Guo, Q., Lian, C., Ren, X., Liang, S., Yu, J., Niu, L., Sun, W., Shen, D. (2019). Automated detection and classification of thyroid nodules in ultrasound images using clinical-knowledge-guided convolutional neural networks. *Medical Image Analysis*, 58: 101555. <https://doi.org/10.1016/j.media.2019.101555>
- [21] Ding, J., Huang, Z., Shi, M., Ning, C. (2019). Automatic thyroid ultrasound image segmentation based on u-shaped network. In *2019 12th International Congress on Image and Signal Processing, Bio Medical Engineering and Informatics (CISP-BMEI)*, pp. 1-5. <https://doi.org/10.1109/CISP-BMEI48845.2019.8966062>

[22] Beyyala, A., Priya, R., Choudary, S.R., Bhavani, R. (2023). Thyroid nodule classification of ultrasound image by convolutional neural network. In: Abraham, A., Bajaj, A., Gandhi, N., Madureira, A.M., Kahraman, C.

(eds) Innovations in Bio-Inspired Computing and Applications. IBICA 2022. Lecture Notes in Networks and Systems, vol 649. Springer, Cham. https://doi.org/10.1007/978-3-031-27499-2_85