SIIM21 ANNUAL MEETING + INDUSTRY PARTNERS
May 24 - 27 | Virtual
Connecting without Boundaries

Automated Risk Stratification of Thyroid Nodules with Ultrasound Cine-clip Images

Rikiya Yamashita, MD, PhD, Stanford University School of Medicine

---

Introduction

Automated risk stratification system for the management of thyroid nodules typically relies on static ultrasound images. However, ultrasound cine images, widely used in clinical practice, are inherently more information-rich and can capture the entire picture of the nodules of interest. This study aims to develop an automated algorithm that uses thyroid ultrasound cine images to classify benign and malignant thyroid nodules.

Hypothesis

We hypothesize our automated risk stratification system could outperform the American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TI-RADS) for classifying benign and malignant thyroid nodules.

Methods

In this IRB-approved study, we retrospectively collected thyroid nodules, which were examined by ultrasound cine images and were subsequently biopsied or resected within a three-month window from the preceding ultrasound scan at our institution from April 2017 through May 2018. We excluded nodules that lack ACR TI-RADS descriptors in the corresponding ultrasound reports. We consider the histopathological diagnosis in the pathology reports as ground truth. Four radiologists performed semi-automated nodule segmentation on the cine-clip images to extract regions of interest (ROI). Our approach learns to extract features using a convolutional neural network from the ROI of each frame of the cine images. To better abstract the feature vectors from cine images, we employ a Transformer architecture, which is designed to handle sequential data, such as natural language, to capture the local and global contexts by using an attention mechanism. We applied five-fold cross-validation for model performance evaluation. The original class distribution was maintained in each fold. We assessed the model performance using the area under the receiver operating characteristic curve (AUROC). Statistical comparison was performed using a paired t-test for average AUROC derived from cross-validation. A two-tailed alpha criterion of 0.05 was used for statistical significance.

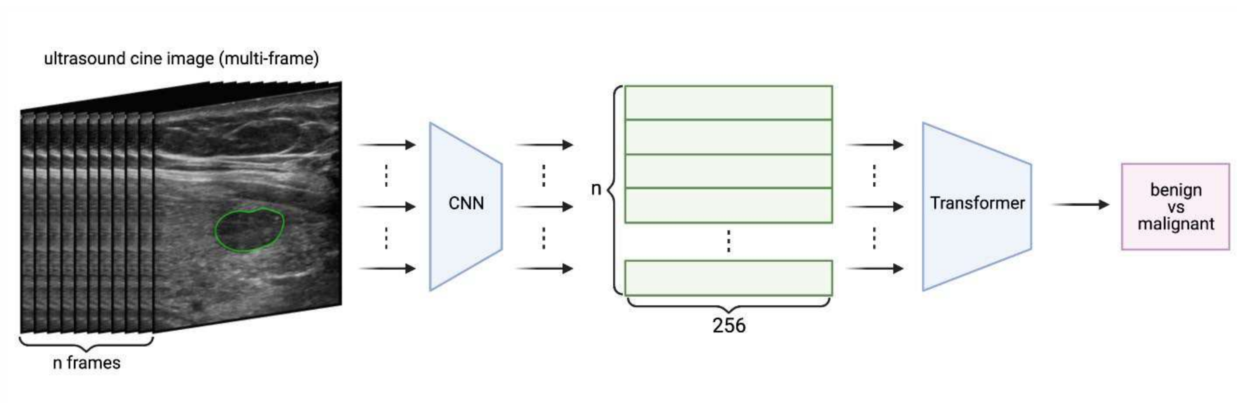
Results

Our dataset comprises 192 thyroid nodules (175 benign and 17 malignant) from 177 patients. The total number of the image frames was 17,412, where the mean, standard deviation, and range per nodule were 90.7, 56.1, and [11, 442], respectively. Our automated system achieved an average AUROC of 0.858 (standard deviation; 0.102) and outperformed the classification performance of the ACR TI-RADS of 0.798 (standard deviation; 0.009); however, it failed to demonstrate a statistically significant difference (t-statistic=1.280, p-value=0.270).

Conclusion

Using ultrasound cine images, our automated risk stratification system outperformed the ARC TI-RADS in classifying benign and malignant nodules, though no statistically significant difference was presented.

Figure(s)



Overview of our system using thyroid ultrasound cine images

Keywords

Artificial Intelligence; Imaging Research