

# Ultrasonographic Thyroid Nodule Classification Using a Deep Convolutional Neural Network

Soon Woo Kwon, BA<sup>1\*</sup>, Ik Joon Choi, MD, PhD<sup>2\*</sup>, Won Il Jang, MD, PhD<sup>3</sup>, Guk-Haeng Lee, MD, PhD<sup>2</sup>, Myung-Chul Lee, MD, PhD<sup>2</sup>

<sup>1</sup>Radiation Medicine Clinical Research Division, Korea Institute of Radiological and Medical Sciences (KIRAMS), Seoul, Korea

<sup>2</sup>Department of Otorhinolaryngology, <sup>3</sup>Radiation Oncology, Korea Cancer Center Hospital, Korea Institute of Radiological and Medical Sciences (KIRAMS), Seoul, Korea



## Background

The objective of this study is to develop a thyroid nodule classification system based on ultrasonography using a DCNN. To evaluate the proposed system, its results are compared with those of previous methods.

## Methods

In this study, transverse and longitudinal ultrasonographic thyroid images of 762 patients from the Korea Institute of Radiological & Medical Sciences were used for training and testing a deep learning model. Of all the cases, 325 were confirmed as benign (nodular hyperplasia, follicular adenoma, or cyst), and 437 were confirmed as papillary thyroid carcinoma after surgical biopsy. Each image has only one thyroid nodule, and an expert physician drew a rectangular region of interest around the pathologic nodule. The images were extracted from thyroid ultrasound video sequences captured with EPIQ 5G, HI VISION Ascendus and EUB-7500 ultrasound devices. The images were in JPEG format and ranged from 640 × 480 to 1024 × 768 pixels in size. All devices employed 12 MHz convex and linear transducer settings. The extracted images were distributed into training, validation, and test sets in a ratio of 6:2:2, as summarized in Table 1.

Table 1. Distribution of cases in training, validation, and test datasets

	Benign			Cancer				
	Total	A	B	C	Total	A	B	C
<b>Training</b>	199	108	52	39	260	101	86	73
<b>Validation</b>	64	39	13	12	94	35	34	25
<b>Test</b>	62	33	15	14	83	34	26	23

## Image Preprocessing

To remove annotations such as the caliper marks used to locate and measure nodule size as well as restore the gaps with the textures surrounding the annotations after removal, the input ultrasonographic images were processed as follows (Fig. 1A, B). The Roberts cross operator implemented in the scikit-image library, which is a Python image processing module, was used to detect edges. Then, pixels greater than 0.25 in value were identified as artifacts. The pixel values of the detected artifacts were deleted in the original image and the missing regions were recovered using the “inpaint\_biharmonic” function from the scikit-image library with the default parameters.

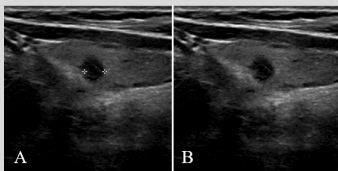


Fig 1A, B

## Sample Augmentation

Nodule-centered images were cropped to the largest rectangular shape in each of the original transverse and longitudinal sonographic images. Two random-sized rectangular images including a nodule were cropped from each vertex. Thus, eight rectangular images including a nodule and one vertex were cropped from the original images (Fig. 2).

In addition, considering the bilateral symmetry of the thyroid, horizontal flip was applied to these images. In summary, 18 transverse and 18 horizontal images per patient were created from the largest rectangular image plus the eight smaller rectangular images.

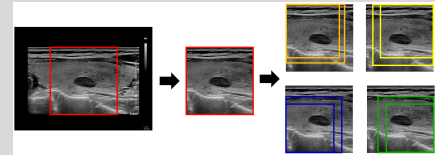


Fig 2

## Transfer learning by fine-tuning modified VGG16 model

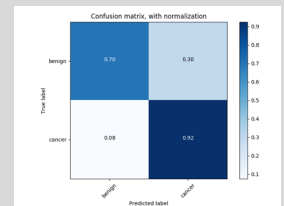
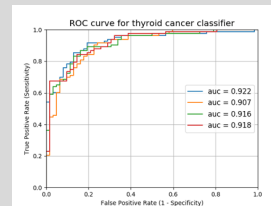
Because of the insufficient number of training samples, we employed a transfer learning method with a pretrained deep learning model. The model used in this study was VGG16, which is known for its good classification results in the ImageNet Large Scale Visual Recognition Challenge. We modified VGG16 by replacing the fully connected layer with global average pooling and sigmoid layers (Fig. 3). The binary cross entropy function was used for the loss function, and the network was minimized by an Adam optimizer at an initial learning rate of  $4 \times 10^{-6}$  with a decay rate of  $10^{-6}$ . The batch size was set to three. All layers in the modified VGG16 were fine-tuned. The Keras (version 2.1.5) wrapper deep learning library with TensorFlow (version 1.7) was used as a backend with Python version 3.6.5 for implementing the modified VGG16 deep learning model. Our calculation was performed on a computer running 64-bit Windows 10 and equipped with one Geforce 1080 Ti. All of the input images were rescaled to [0,1] and then the mean of each image was set to zero using a Keras built-in operation. Four-fold cross-validation was performed to prevent overfitting. The weights of the fine-tuned deep learning model were obtained at the lowest validation loss. Figure 4 shows the detailed process of fine-tuning the VGG16 model in this study.



Fig 3

## Results

Area Under the Curve	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0.916 [0.907 – 0.922]	0.92	0.70	0.90	0.75



## Conclusion

In conclusion, we introduced a new fine-tuned deep transfer learning model for classifying thyroid nodules in ultrasonography. We expect this model will help physicians diagnose thyroid nodules with ultrasonography.