THYROID NODULE CLASSIFICATION IN ULTRASOUND IMAGES: A DEEP LEARNING APPROACH USING SRI LANKAN DATA

Rinusga R. [1,*]

[1] Department of Physical Science, Faculty of Applied Science, University of Vavuniya, Sri Lanka.
[*] Corresponding author email: rrinusga@gmail.com

Abstract: Thyroid nodules have garnered worldwide attention due to thyroid cancer ranking as the fifth most prevalent cancer among women. In Sri Lanka, the concern is even more pronounced, as it stands as the second most common cancer in females and the primary cancer affecting those aged 15-34. The diagnostic process involves a comprehensive assessment, encompassing medical history, physical examinations, thyroid function tests, imaging, and occasionally, biopsy. Among these, high-resolution ultrasound stands out as the gold standard for non-invasive, real-time, and cost-effective nodule identification. In this study, we introduce a novel deep learning approach, meticulously tailored to the distinctive characteristics of Sri Lanka. This endeavor represents the pioneering use of deep learning techniques with ultrasound images from Sri Lankan subjects. Our objectives encompass the development of a comprehensive thyroid ultrasound image classification system, the creation of a specialized database, and the design of an efficient deep learning architecture. The data used for this study was collected from the National Hospital of Sri Lanka in 2023, encompassing 1,458 ultrasound images derived from 195 patients. These images are annotated with Thyroid Imaging Reporting and Data System (TI-RADS) classes, resulting in two datasets: Dataset 01 (Benign or Malignant) and Dataset 02 (Benign, Mildly Suspicious, or Malignant). Preprocessing steps were undertaken to eliminate extraneous data and incorporate data augmentation techniques. The thyroid nodule datasets were then divided into training and testing sets, with an 80% allocation for training and 20% for testing. Our approach involved training three distinct deep learning classification models based on VGG-16, VGG-19, and ResNet-50, employing a rigorous 5-fold cross-validation methodology. The results showed that VGG-19 achieved an impressive accuracy of 88% for Dataset 01, while ResNet-50 achieved a solid 80% accuracy for Dataset 02. This utilization of deep learning models, in conjunction with pre-trained networks and the use of specialized Sri Lankan data, has demonstrated remarkable accuracy and clinical relevance. This research holds the promise of revolutionizing early detection and refining the diagnosis of thyroid nodules, offering the potential to enhance healthcare outcomes not only in Sri Lanka but also beyond.

Keywords: Classification, Deep learning, Sri Lanka, Thyroid nodules, Ultrasound images