Analysis of Carpal Tunnel Syndrome: Simultaneous Ultrasound Median Nerve Segmentation and CTS Classification Enabled by Deep Learning

Hong-kun Lin [1] Ting-Hsuan Chiu[2] Hao-Jen Wang[3] Chia-Yen-Lee[4*]

[1]Department of Biomedical Engineering, National Yang Ming Chiao Tung University, Taiwan

[2]Department of Biomedical Engineering & Environmental Sciences, National Tsing Hua University, Taiwan

[3]Department of Biomedical Engineering, National Taiwan University, Taiwan

[4]Department of Electrical Engineering, National United University, Taiwan

I. Abstract

Carpal tunnel syndrome (CTS) is commonly caused by prolonged hand movements, leading to compression of the carpal tunnel, which affects the median nerve, resulting in numbness and pain. Ultrasound diagnosis, known for its non-invasive and ease of use, is often employed to assist in CTS diagnosis. However, when doctors manually annotate multiple images to observe the dynamic state of the median nerve, the presence of speckle noise in ultrasound images can affect diagnosis and image segmentation. Therefore, this study developed a model to diagnose CTS and simultaneously provide median nerve annotations. The model achieved a diagnostic performance with an accuracy of 0.83, sensitivity of 0.77, specificity of 0.9, precision of 0.91, and an F1 score of 0.83. In terms of median nerve location annotation, the model achieved a Dice score of 0.81 ± 0.22. Additionally, the study utilized YOLOv7, a high-performing model in various image classification and segmentation tasks, as a comparison, and the results showed that the proposed method significantly outperformed YOLOv7.

II. Materials and methods

Materials

- The ultrasound image sequences were provided by Kuang Tien General Hospital in Taichung.
- The dataset consists of 2623 ultrasound images, including 11 healthy individuals and 10 CTS-diagnosed patients confirmed through nerve conduction studies.
- The training data includes 980 images of healthy individuals and 828 images of CTS patients; the validation data includes 179 healthy images and 91 CTS images; and the test data includes 269 healthy images and 367 CTS images.

Methods

The study process is divided into two parts. The first part involves image preprocessing, where bias field correction is applied to remove speckle noise from ultrasound images, separating the median nerve from the image background. Subsequently, Multi-Frame Concatenation is employed to combine information from previous and subsequent frames, mimicking the behavior of physicians who refer to adjacent frames when the median nerve is not clearly visible. The second part focuses on model development. Mask R-CNN is used as the baseline model, integrated with the Convolutional Block Attention Module (CBAM). By incorporating attention mechanisms, the model is able to focus on the median nerve in the images.

Bias Field Correction(BC)

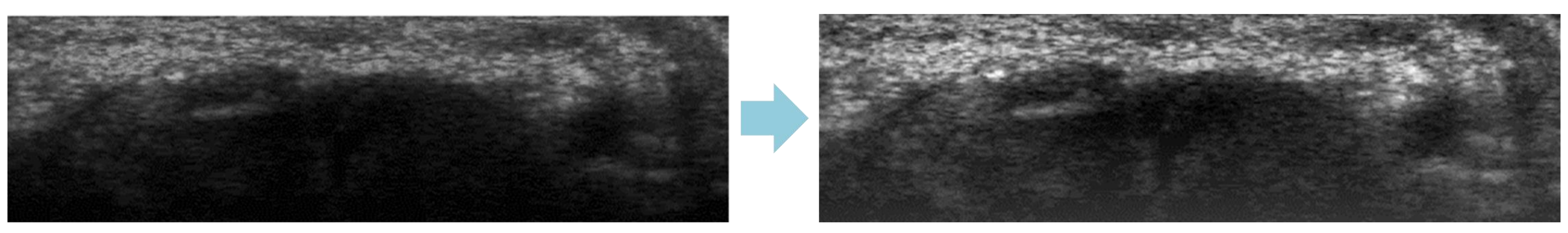

Figure1:Bias field correction

Multi-Frame Concatenation(MFC)

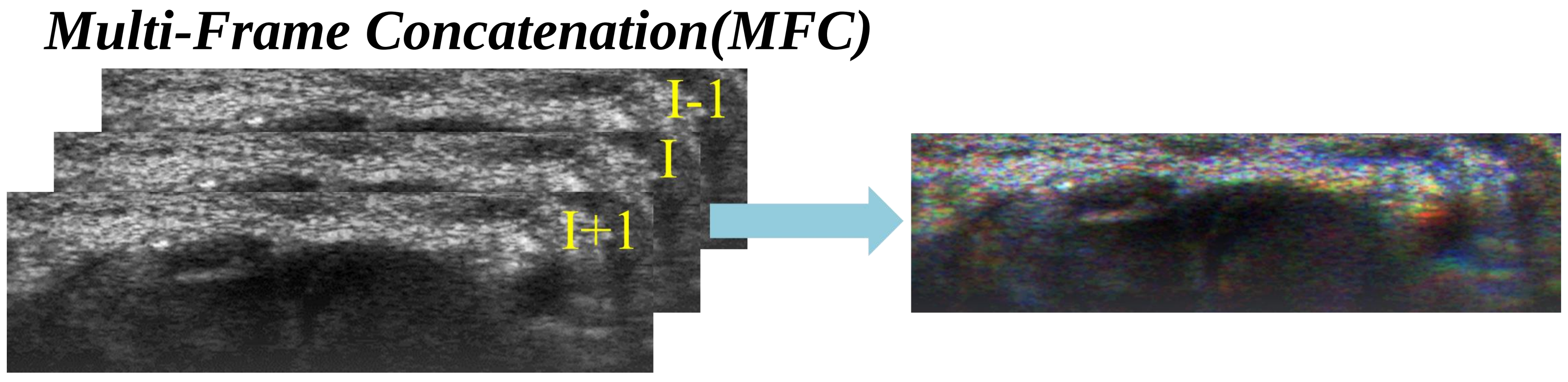

Figure2:Multi-Frame Concatenation

Convolutional Block Attention Module(CBAM)

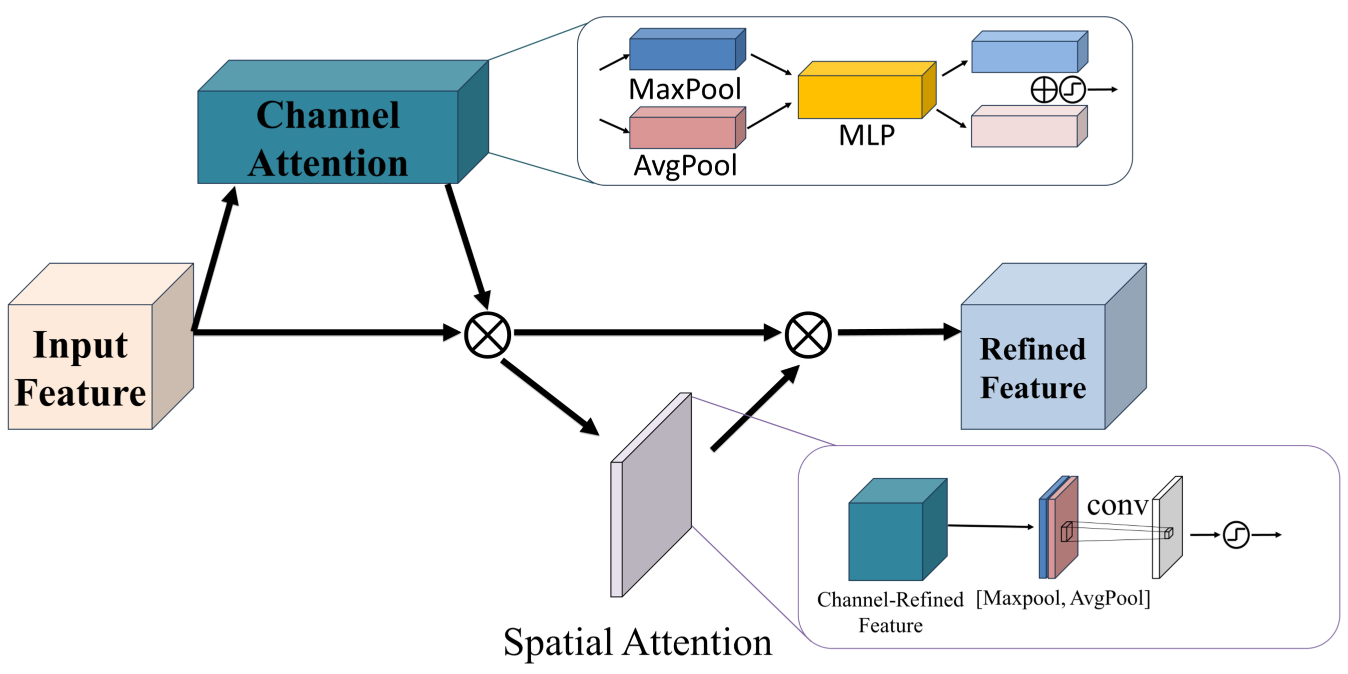

Figure3:CBAM

Research flow chart

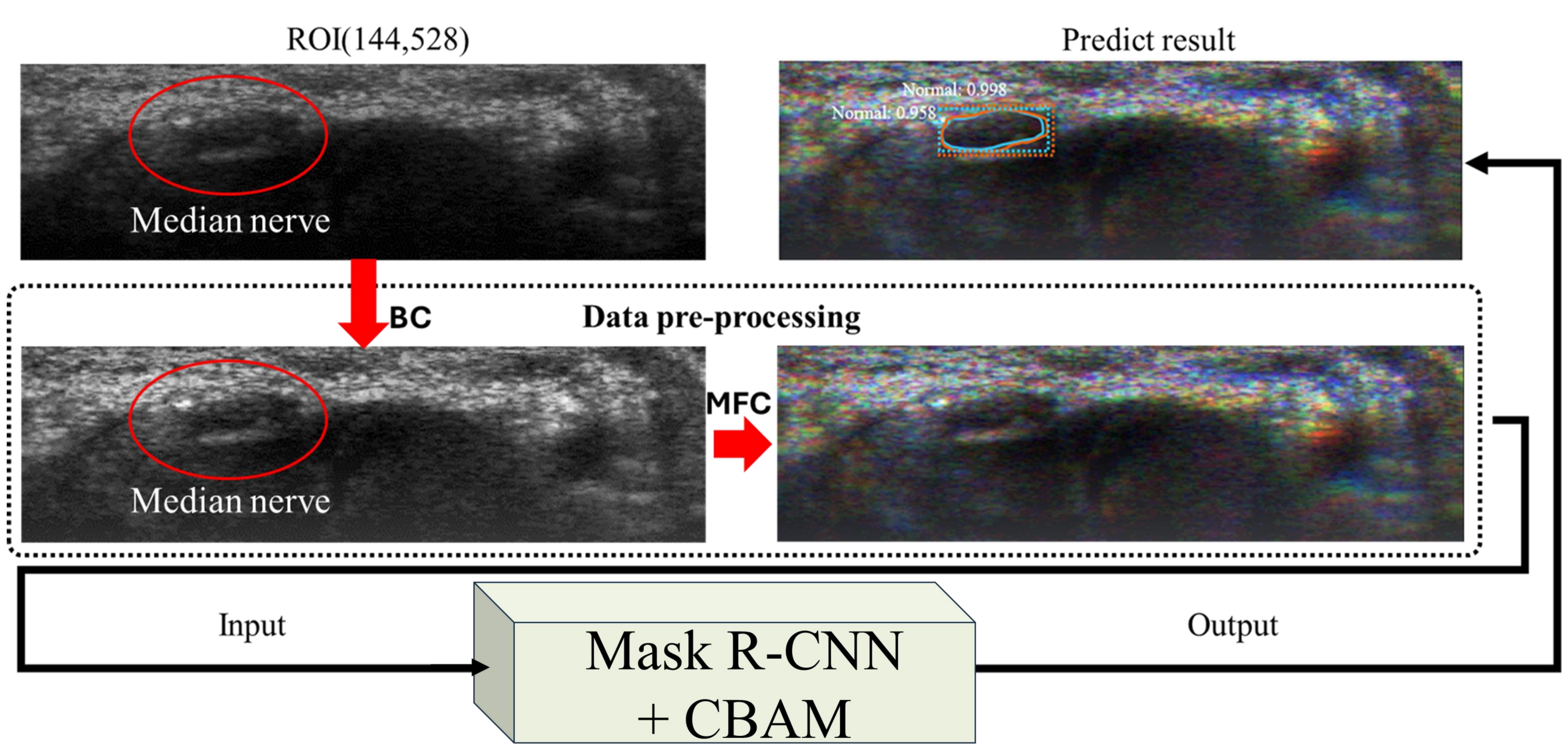

Figure4:Proposed model

III. Results

- This study compares the performance of Mask R-CNN combined with CBAM in predicting CTS and annotating the median nerve.
- As shown in Table 1 and Table 2, the combination of Mask R-CNN and CBAM significantly improves the accuracy, sensitivity, and f1 score for predicting CTS, and notably enhances the dice performance in the median nerve annotation task.
- This study used YOLOv7 for comparison, and the results showed that combining Mask R-CNN with CBAM significantly outperformed YOLOv7.

Table 1. CTS Classification Performance

Classification	*Accuracy*	*Sensitivity*	*Specificity*	*Precision*	*F1 score*
Mask R-CNN + CBAM	**0.83**	**0.77**	**0.9**	**0.91**	**0.83**
Mask R-CNN	**0.68**	**0.45**	**1.0**	**0.99**	**0.62**
YOLOv7	**0.68**	**0.44**	**0.96**	**0.93**	**0.60**

Table 2. Median Nerve Segmentation Performance

Segmentation	*Dice*
Mask R-CNN + CBAM	**0.83**
Mask R-CNN	**0.68**
YOLOv7	**0.68**

IV. Conclusions

- To accurately, effectively, and objectively assist doctors in diagnosing CTS, this study proposes a model for CTS diagnosis and annotation.
- In the preprocessing of ultrasound images, boundary correction is used to resolve the issue of blurred boundaries of the median nerve, and multi-frame concatenation is applied by referencing information from preceding and succeeding frames to help the model better locate the median nerve.
- The proposed model achieved performance in CTS classification with an accuracy of 0.83, sensitivity of 0.77, specificity of 0.9, precision of 0.91, and an F1 score of 0.83. For the median nerve segmentation task, the model achieved a Dice score of 0.8.
- This study used YOLOv7, a model known for its excellent performance in various image classification and segmentation tasks, as a comparison. The results showed that the proposed method significantly outperformed YOLOv7 in both median nerve segmentation and CTS classification tasks.