노인재활

발표일시 및 장소: 10 월 19 일(토) 14:10-14:20 Room B(5F)

OP2-4-2

Auto-segmentation of vocal fold through convolution neural network using VFSS

Hyun Haeng Lee[1*], Bomi Kwon[1], Joe Yeh[2], Jongmin Lee[1†]

Konkuk University Medical Center, Department of Rehabilitation Medicine[1], aetherAI, aetherAI[2]

Objective

The movement of the vocal fold can be used as an indirect indicator of the degree of laryngeal elevation during swallowing reflex. In this study, we aimed to demonstrate the feasibility of autosegmentation for vocal fold in video of videofluoroscopic swallowing study (VFSS) through convolution neural network, which is expected to be used in automatic measuring system for laryngeal elevation.

Methods

We selected 12 videos of VFSS randomly among VFSS performed at Department of Rehabilitation Medicine, Konkuk University Medical Center from 2010 to 2017. A designated skillful researcher drew outline of spinous process of first cervical vertebra, vocal fold, oropharynx, laryngopharynx, and larynx on each frame of VFSS video (Fig.1). The outline drawn frames were divided into training dataset and test dataset, which were used for constructing and validating the autosegmentation model by convolution neural network (CNN). Outline drawn by the researcher was used as label. The details of the CNN framework are shown in Fig.2. We modified each frame of the video by random rotation (-30 to 30 degree), random scale (0.8 to 1.2), and random shift (-10 to 10%) to ensure external validity. We used Gaussian blur to reduce image noise and detail. The outcome was presented as intersection over union (IOU), which is calculated as the area of overlap over the union. Area of overlap refers to the overlapped area of autosegmentation obtained through CNN and the outline drawn manually by the researcher. The process of model construction and validation through CNN was conducted through collaboration with aetherAI, Taiwanese artificial intelligence company.

Results

The designated researcher drew the outline of spinous process of first cervical vertebra, vocal fold, oropharynx, laryngopharynx, and larynx in 5620 frames of 12 VFSS videos. The test data and the training data were divided at the ratio of 11:1. The test results for the model constructed by CNN were 0.47 IOU for vocal fold and 0.73 IOU for spinous process (Fig.3).

Acknowledgment: This work was supported by Konkuk University Medical Center Research Grant 2019

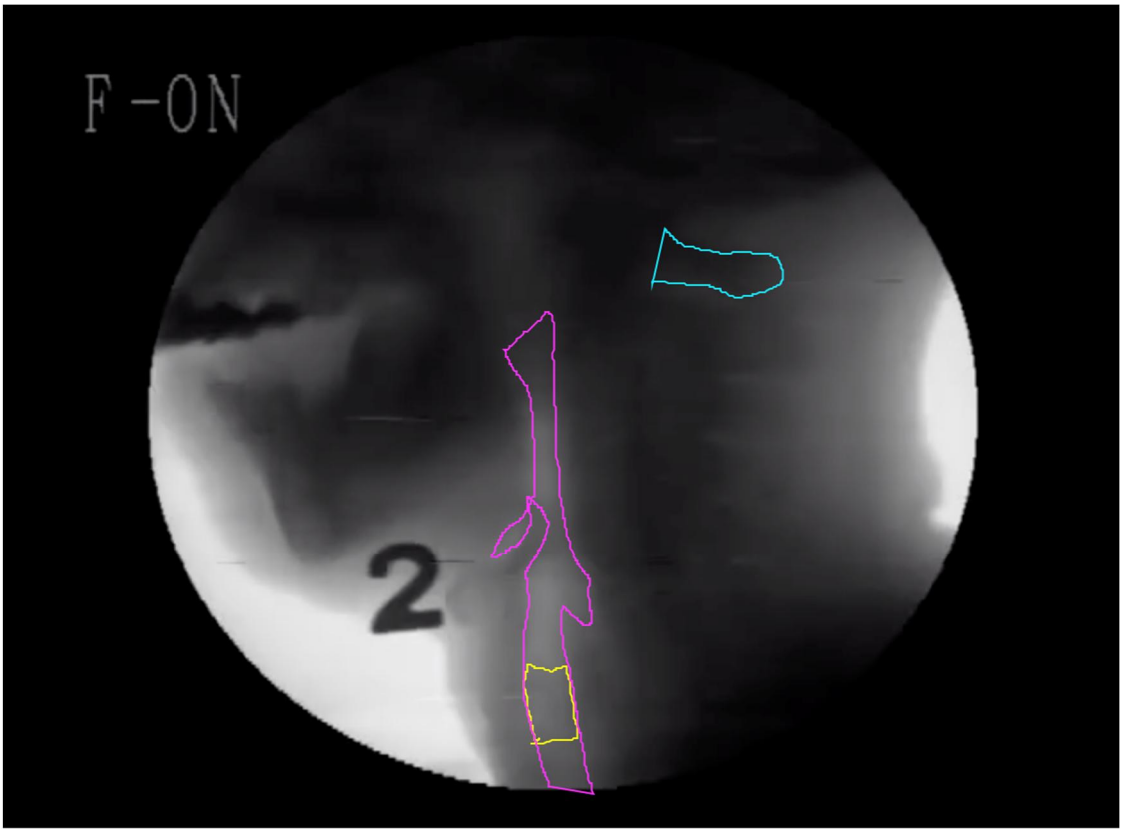

Fig.1. Drawn outline of spinous process of first cervical vertebra, vocal fold, oropharynx, laryngopharynx, and larynx

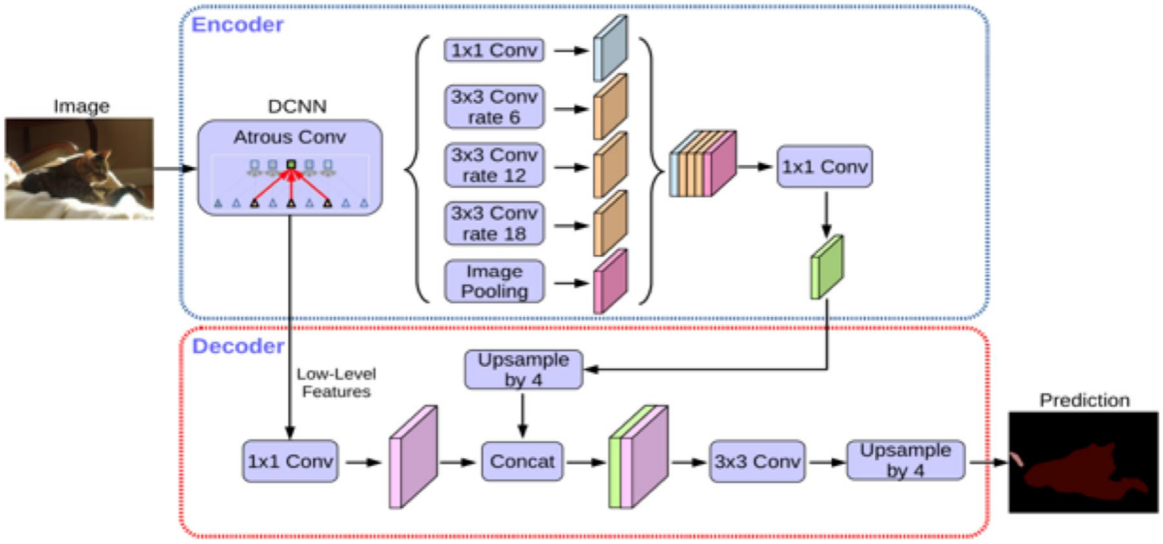

Fig.2. Framework of convolution neural network

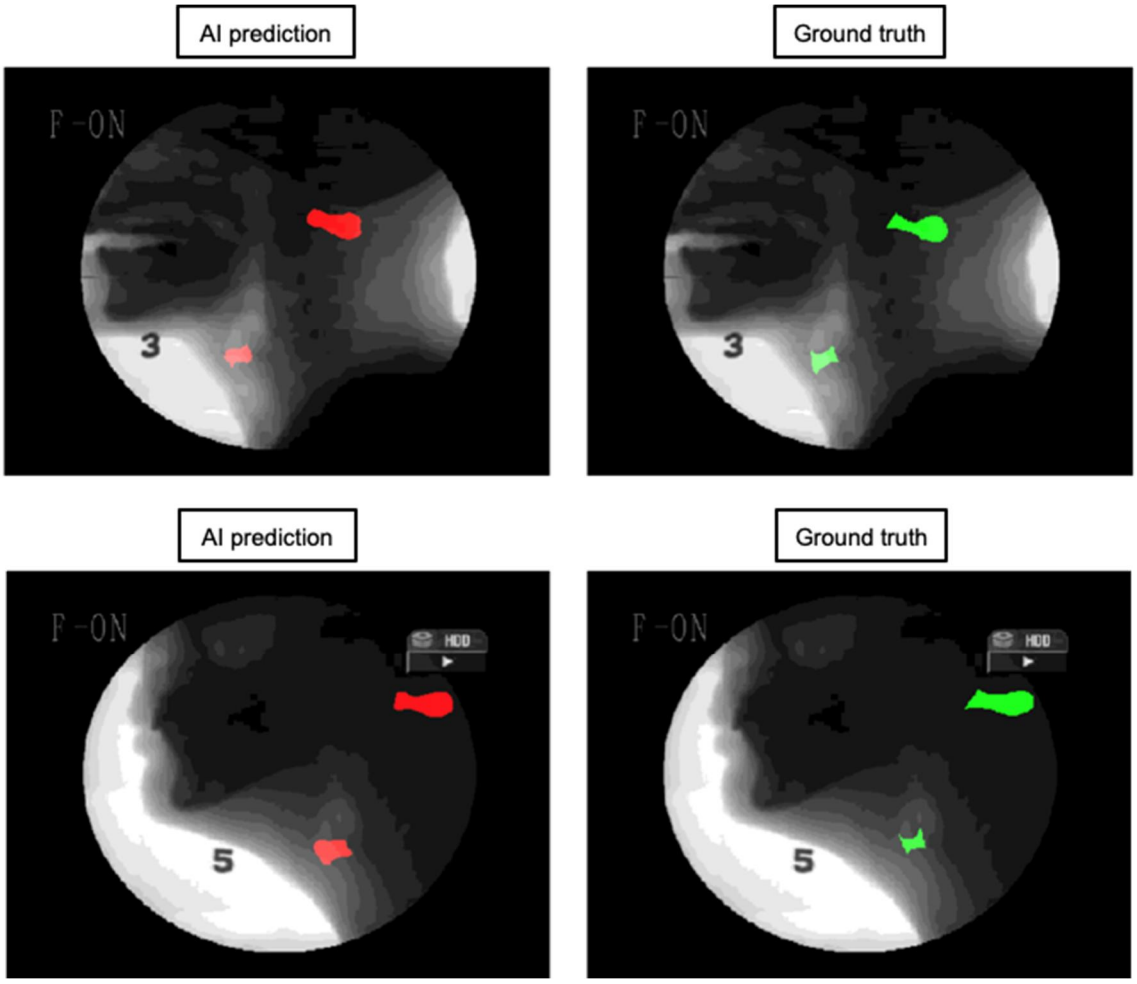

Fig.3. Comparison of auto-segmentation by CNN and outline drawn by human researcher