P 3-9

Detection of Pharyngeal Phases in the Videofluoroscopic Swallowing Study using Deep Learning

Eunhee Park[1,3*], Yu-Sun Min[1,3], Jae-Won Huh[2,3], Won-Jong Yang[1,3], Tae-Du Jung[1,3†]

Kyungpook National University Chilgok Hospital, Department of Rehabilitation Medicine[1], Kyungpook National University Hospital, Department of Rehabilitation Medicine[2], School of Medicine, Kyungpook National University, Department of Rehabilitation Medicine[3]

Introduction

Previous computer assisted analysis of videofluoroscopic swallowing study (VFSS) required manual annotation tools to enter several defined anatomical positions and selected time intervals of interest. These processes were still costly and challenging for clinicians.

Objective

In this study, we present a novel approach to detect a pharyngeal phase of swallowing through whole of VFSS video clips using inflated 3D Convolutional Networks (I3D) without additional manual annotations.

Methods

The VFSS data were collected from 144 subjects (44 females, aged 63.3±16.4 years) who complained subjective difficulties of swallowing during diets and visited the inpatient and outpatient clinic in Department of Rehabilitation Medicine. We propose a two-stage process to detect pharyngeal phases in the VFSS video instances. First, we detect intervals of frames having maximal vertical movements. Next stage, we trained I3D networks to classify the interval of video frames with the label whether the interval is in a pharyngeal phase or not. We evaluated I3D networks of single stream such as optical flow (I3D-Flow), RGB (I3D-RGB), and two streams (I3D-Joint).

Results

When the I3D-Flow and the I3D-RBG networks were trained by 10K iterations, the network showed 93.94% and 95.91% accuracy rate, separately. When the I3D-Joint was trained by 30K iterations, the network showed 95.64% accuracy rate which is similar to I3D-RGB.

Conclusion

We have presented a novel framework based on I3D networks to detect a pharyngeal phase during swallowing in whole VFSS video clips. The algorithm is validated on large clinical dataset and achieves the state-of-the-art of human swallowing motion analysis. This study showed that a new framework could classify the pharyngeal phase of swallowing without any manual adjustments and should be a fundamental work for developing software of fully automatic analysis in VFSS.

Acknowledgment

This work was supported by Biomedical Research Institute grant, Kyungpook National University Hospital(2018)