뇌신경재활

발표일시 및 장소: 10 월 19 일(토) 10:20-10:30 Room B(5F)

OP2-3-3

Automatic Detection of Penetration or Aspiration in Videofluoroscopy by Deep Learning Technology

Hyun Il Kim[5], Tae Uk Kim[1], Seo Young Kim[1], Jung Keun Hyun[1,2], Sang Il Choi[4], Seong Jae Lee[1†], Joo Young Ko[1*]

Dankook University Hospital, Department of Rehabilitation Medicine[1], Dankook University, Department of Nanobiomedical Science & BK21 PLUS NBM Research Center for Regenerative Medicine[2], Dankook University,) Institute of Tissue Regeneration Engineering (ITREN)[3], Dankook University, Department of Applied Computer Engineering[4], Dankook University, Department of Software Engineering[5]

Objective

Analysis of Videofluoroscopic swallowing study (VFSS) can be time-consuming and tedious job and its reliability has been reported to be unsatisfactory. Recent advancement in image classification technology enables fully automatic detection of objects but its application but it has never been applied in VFSS image analysis. We aimed to develop an algorithm that can detect penetration or aspiration in VFSS images in a fully automated manner, using open source deep learning algorithms including convolutional neural network. This study may be the first step of development of fully automatic reading system of VFSS.

Methods

One hundred eighty VFSS files (108 normal and 72 with penetration or aspiration) from 71 subjects were collected for establishment of dataset. Each video files were split into individual frame images obtaining 138,890 normal and 5,682 dysphagic images with penetration or aspiration. Determination of normal and dysphagic images was performed by a physiatrist experienced in VFSS. Dataset was balanced by undersampling of normal images to 23,465 images and oversampling of dysphagic images to 24,672 images. Images were distributed to training, validation and test images at the ratio of 7:1:2. After equalizing by contrast limited adaptive histogram equalization (CLAHE) algorithm, automated detection of penetration or aspiration was trained by open source deep learning algorithms. Detection was accomplished in two steps. First, regions of interest (ROIs) were set around larynx with reference to cervical spinal column identified by cervical detector using Unet deep learning segmentation model that showed accuracy of 99.4% in previous study. Then penetration/aspiration detector identified presence of penetration or aspiration in those ROIs. Accuracy based on frame or video file was calculated and compared among the deep learning models.

Results

Accuracy of each deep learning model is presented in tables. Xception model showed the best result in frame based analysis but NasnetMobile did in file based analysis. Accuracy of

detecting frame images that showed penetration or aspiration was 98.6%. When classifying files with penetration or aspiration, accuracy was 89.7%.

Conclusion

This study was first attempt to identify penetration or aspiration from VFSS images in a fully automated manner using deep learning image classification technology. The results show that deep learning algorithm can detect penetration or aspiration automatically from VFSS images with a significant accuracy. Further researches to improve accuracy and prospective clinical trials are required.

Acknowledgment: 1. This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (grant number: 2018R1D1A3B07049300) 2. This research was supported by the MISP(Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW (2017-0-00091), supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)

Table1. Frame based accuracy of detection by deep learning model

Deep Learning Algorithm	Accuracy[1)]	Recall[2)]	Precision[3)]	Specificity[4)]
Resnet50	94.3%	92.1%	43.9%	94.4%
NasnetMobile	96.8%	90.1%	59.5%	97.1%
Xception	98.6%	91.4%	81.0%	99.0%

1) Accuracy = $\frac{\text{Number of images (normal and dysphagic) correctly predicted by deep learning model}}{\text{Number of all images}}$

2) Recall (sensitivity) = $\frac{\text{Number of images correctly predicted as dysphagic by deep learning model}}{\text{Number of images diagnosed as dyphagic by physiatrist}}$

3) Precision (positive predictive value)=$\frac{\text{Number of images correctly predicted as dysphagic by deep learning model}}{\text{Number of all images predicted as dysphagic by deep learning model}}$

4): Specificity (true negative rate) = $\frac{\text{Number of images correctly predicted as normal by deep learning model}}{\text{Number of images diagnosed as normal by physiatrist}}$

Table2. File based accuracy of detection by deep learning model

Deep Learning Algorithm	Accuracy[1)]	Recall[2)]	Precision[3)]	Specificity[4)]
NasnetMobile	89.7%	89.4%	89.4%	90%
Xception	87.1%	89.4%	85%	85%

1) Accuracy = $\frac{\text{Number of files (normal and dysphagic) correctly predicted by deep learning model}}{\text{Number of all files}}$

2) Recall (sensitivity) = $\frac{\text{Number of files correctly predicted as dysphagic by deep learning model}}{\text{Number of files diagnosed as dyphagic by physiatrist}}$

3) Precision (positive predictive value) = $\frac{\text{Number of files correctly predicted as dysphagic by deep learning model}}{\text{Number of all files predicted as dysphagic by deep learning model}}$

4): Specificity (true negative rate) = $\frac{\text{Number of files correctly predicted as normal by deep learning model}}{\text{Number of files diagnosed as normal by physiatrist}}$