의학 석사학위논문

Deep learning-based interpretation of cerebrovascular reserve on basal/acetazolamide stress brain perfusion SPECT

딥러닝을 이용한 기저/아세타졸아미드 부하 뇌혈류 SPECT에서 뇌혈류 예비능 평가

2018년 8월

서울대학교 융합과학기술대학원

분자의학 및 바이오제약학과

유 현 지

Abstract

Deep learning-based interpretation of cerebrovascular reserve on basal/acetazolamide stress brain perfusion SPECT

Hyun Gee Ryoo

Department of Molecular Medicine and Biopharmaceutical Science,

The Graduate School of Convergence Science and Technology,

Seoul National University

Early and accurate detection of cerebrovascular disease is important for its mortality and brain injury. Basal/acetazolamide stress brain perfusion single photon emission computed tomography (SPECT) is a functional diagnostic imaging tool to detect cerebral perfusion decrease and cerebrovascular reserve. The visual interpretation of brain perfusion SPECT image is a standard practice in the clinical setting, often resulting interobserver variability and inconsistence of diagnosis. In this study, we applied Long Short-Term Memory (LSTM) network and 3D convolutional neural network (CNN) model for the deep learning-based interpretation of the text report

and image of basal/acetazolamide stress brain perfusion SPECT. LSTM network was successfully trained to classify the text report of each image regarding its hemodynamic abnormality. The LSTM model-predicted results were used for the label of a cerebrovascular reserve decrease on basal/acetazolamide stress brain perfusion SPECT images to train 3D CNN model. Our designed 3D CNN model was trainable but did not show outstanding performance to detect the cerebrovascular reserve decrease on basal/acetazolamide stress brain perfusion SPECT images. Our results suggest that 3D CNN is a trainable model on basal/acetazolamide stress brain perfusion SPECT in the detection of a cerebrovascular reserve decrease using text report prediction of LSTM as a ground truth label. Additional image preprocessing steps with advanced network architecture are required to improve the performance of our deep-learning based interpretation system in future study.

Keywords : deep learning; brain perfusion SPECT; cerebrovascular reserve; Long Short-Term Memory; 3D convolutional neural network

Student Number : 2016-24900

Contents

List of tables

List of figures

Introduction

Cerebrovascular disease is one of the most common cause of death and disability. It is a problem not only in elders but also the young adults and children, and the early detection of disease is important to prevent severe neurological deficit (1-3). For the diagnosis and evaluation of the cerebrovascular disease, various imaging studies such as brain magnetic resonance imaging (MRI), computed tomography (CT) or cerebral angiography are performed.

The functional brain perfusion single photon emission computed tomography (SPECT) has emerged as a potential diagnostic tool in cerebrovascular diseases with good sensitivity (61~74%) and specificity (88~98%) (4-6). The basal/acetazolamide stress brain perfusion SPECT allows noninvasive study of basal perfusion state and cerebrovascular reserve for the patients with ischemic cerebrovascular disease, Moyamoya disease, and subarachnoid hemorrhage. Cerebral hemodynamic evaluation of the patient is important for treatment planning and prognosis prediction. Thus, a careful image interpretation of existence, location, extent and severity of blood flow and cerebrovascular reserve decrease are important for the patient management.

In the clinical setting, visual interpretation is the standard method to determine the existence of perfusion abnormality in brain perfusion SPECT. However, the visual interpretation is often inconsistent because it causes interobserver variability.

Manual or automatic quantification methods were tried to overcome the weakness, but these are still not practically used (7-9).

The recent dramatical development of deep learning allowed machines to learn representation of large data set with multilevel abstraction (10). A great number of studies are performed to apply this method in the medical field. Especially, deep convolutional network made a breakthrough in image recognition, allowing medical image detection, segmentation, and classification (11-13). To our knowledge, there was no previous study using deep learning architecture for basal/acetazolamide stress brain perfusion SPECT interpretation. In this study, we aimed to develop a deep learning-based interpretation model of cerebrovascular reserve on basal/acetazolamide stress brain perfusion SPECT using text reports and images and evaluated its performance.

Material and methods

Subjects

The data used in this study were obtained from Seoul National University Hospital (SNUH) cohort. Text reports and images of basal/acetazolamide stress brain perfusion SPECT from Dec 2008 to Nov 2017 were retrospectively collected. Patients underwent brain perfusion SPECT for initial diagnosis, medical follow-up, preoperative evaluation, or postoperative evaluation of diseases. All formal reports for each scan were confirmed by more than 2 nuclear medicine specialists.

The retrospective study using data from SNUH cohort was approved by the institutional review boards (IRB) of our institute and the informed consent was waived due to the retrospective design. All procedure performed in this study involving human participants were accordance with ethical standards of the institutional and with the 1964 Helsinki declaration and its lateral amendments or comparable ethical standards.

Basal/acetazolamide stress brain perfusion SPECT images

SPECT images were acquired by dedicated triple-head gamma camera (TRIONIX Triad XLT 3, Trionix Research Laboratory, Inc., Twinsburg, OH, USA) with Fan-Beam collimator. For the basal/acetazolamide stress brain perfusion SPECT study,

740 MBq and 1480 MBq of ^{99m}Tc-hexamethylpropyleneamine oxime (HMPAO) was intravenously injected and each image was acquired 2-3 minutes after injection. The acetazolamide was intravenously injected 15 minutes before the acetazolamide stress scan. Images were acquired by protocols of 40 step-and-shoot for 3-degree angle with 22 seconds per each step. Images were reconstructed with as follows: 1) 128 × 128 matrices, 2) filtered back projection, 3) Butterworth filter with high cut frequency of 0.4 and roll off degree of 5.0, 4) Chang's method for attenuation correction. Only the transaxial images of basal/acetazolamide stress brain perfusion SPECT were used to train deep learning model. There was no additional pre-processing step.

LSTM network architecture and training for sentence classification

Total 7658 text reports of brain basal/acetazolamide stress perfusion SPECT were retrospectively collected in SNUH cohort. Only the conclusion of the document was selected for the input data of long short-term memory (LSTM) network model. All the letters in the document were written in English character. Among the 7658 reports, 293 reports were randomly selected and one nuclear medicine physician interpreted and labeled the existence of basal perfusion and vascular reserve decrease in each given 14 categories as follows: basal perfusion in right anterior cerebral artery (ACA) territory, vascular reserve in right ACA territory, basal perfusion in left ACA territory, vascular reserve in left ACA territory, basal perfusion in right middle cerebral artery

(MCA) territory, vascular reserve in right MCA territory, basal perfusion in left MCA territory, vascular reserve in left MCA territory, basal perfusion in the right hemisphere, vascular reserve in the right hemisphere, basal perfusion in the left hemisphere, vascular reserve in the left hemisphere, basal perfusion in the whole brain, vascular reserve in the whole brain, respectively.

We designed LSTM framework and architecture is summarized in Figure 1. Non-characters in the sentence such as number, commas, periods, and other punctuation was replaced in to whitespace. The words were split into word tokens using whitespaces as separators. The word tokens were transformed into vectors, the process called word embedding. The extracted feature of each word was fed into LSTM network. The manually labeled data were used to train LSTM network for conclusion prediction. The output of the network had two nodes for 14 dense layers, which respectively corresponds to its hemodynamic state whether it is normal or decreased.

The training was conducted by the method for gradient-based optimization algorithm called *Adam* (14). 90.0% of text data (263/293) were used for the training set and the remaining 10% (30/293) were used for the validation to monitor its performance. The LSTM network model was trained for 70 epochs, and the momentum parameter was set to 0.9. The learning rate was initially 0.001 and logarithmically decreased to have $1x10^{-5}$ at the final epoch.

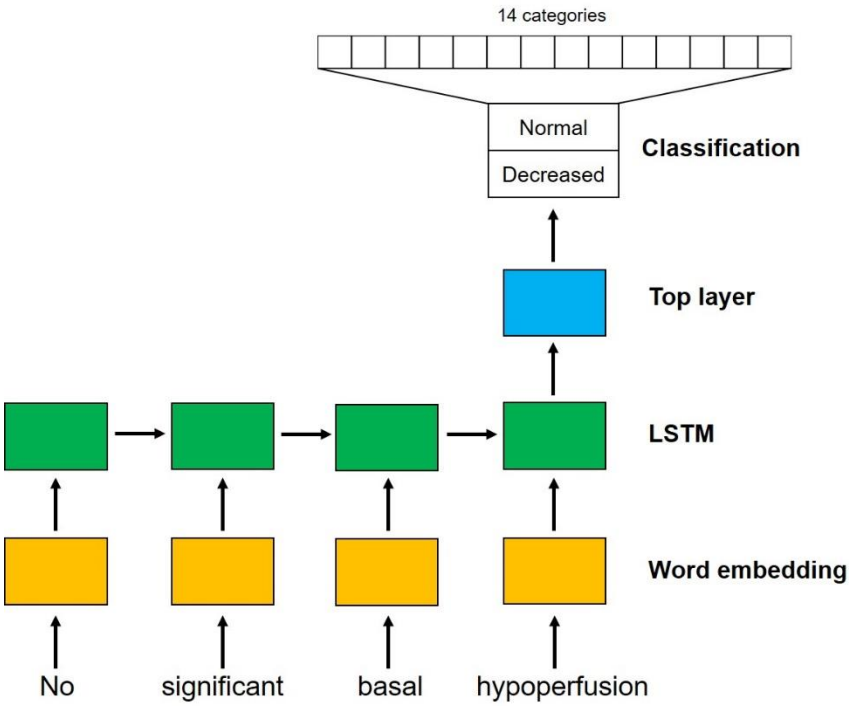


Figure 1. LSTM network architecture for classification of text report of brain perfusion SPECT

Conclusion of the text report of the basal/acetazolamide stress brain perfusion SPECT was used for the input of LSTM network. Each separated word was embedded and fed into LSTM network. Manually labeled data were used to train LSTM network to predict the basal perfusion and cerebrovascular reserve decrease in each given 14 categories.

3D CNN architecture and training for image classification

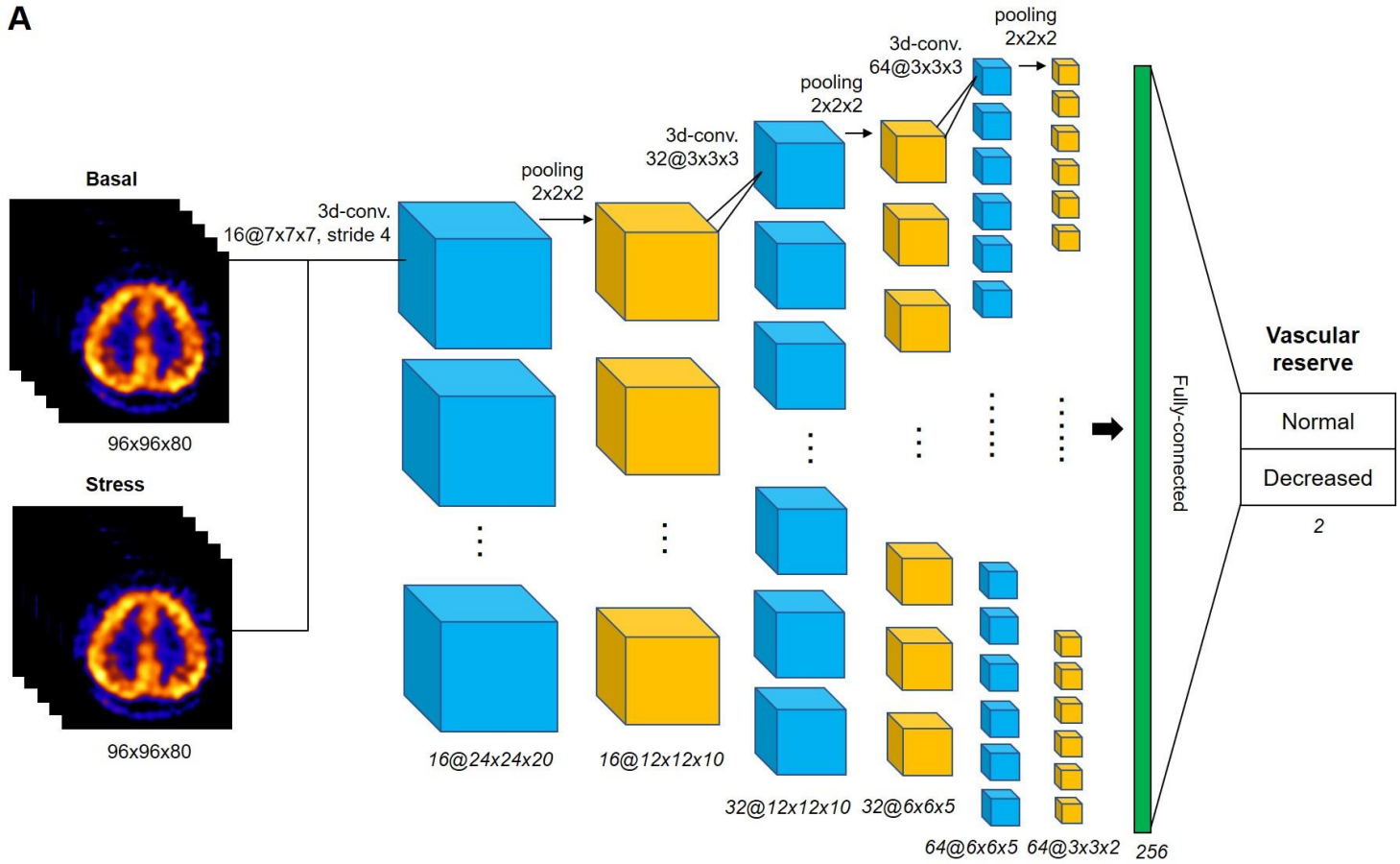
We designed a 3-dimensional convolutional neural network (3D CNN) framework and architecture is summarized in Figure 2a. Total 2018 brain perfusion SPECT images of patients from Jan 2015 to Nov 2017 were used and the demographics of images are summarized in Table 1. Both basal and acetazolamide challenged brain perfusion SPECT images of each patient were used in two different input layers. The LSTM model-predicted vascular reserve of the whole brain was used as a ground truth label of each image set.

The 3D volume of images was reshaped to have 96 x 96 x 80 matrix dimension. If the raw image has smaller dimension, zero-padding along each dimensional axis (x-, y- and z-axes) were applied. If the raw image has greater dimension, the most front and last values were sequentially subtracted. Input values of voxels were rescaled by the range of -1 to 1.

The images were passed by 3 convolution filters with size of (16@7 x 7 x 7), (32@3 x 3 x3) and (64@3 x 3 x 3) in sequence. The stride size of four voxel was applied for the first convolution filter. After each convolutional layer, Rectified Linear Unit (ReLU) activation layer and max-pooling were followed. For max pooling operation, the pooling layer with size of (2 x 2 x 2) were applied respectively. 3-dimensional vector was flattened to 1-dimensional vector to produce fully-connected layer with 256 feature vectors. Two times of drop-out technique were applied after the linear layer and before the activation function layer. A softmax function, exponential activation function with normalized operator, was applied to

discriminate two labels after the output of the fully-connected layer. The network was trained to minimize the cross-entropy loss between the predicted result and the label of images.

The training was conducted by the method for gradient-based optimization algorithm called *Adam* (14). 90.0% of image data (1811/2018) were used for the training set and the remaining 10% (202/2018) were used for the validation to determine architecture and parameters including epoch, number of nodes, layers, and learning rate. Our designed 3D CNN model was trained for 17 epochs, and the momentum parameter was set to 0.9. The learning rate was initially 0.0001 and logarithmically decreased to have 1×10^{-6} at the final epoch. Balancing of training data was done by class weight adjustment.


A
Basal
96x96x80
Stress
96x96x80
3d-conv.
16@7x7x7, stride 4
pooling
2x2x2
3d-conv.
32@3x3x3
pooling
2x2x2
3d-conv.
64@3x3x3
pooling
2x2x2
16@24x24x20
16@12x12x10
32@12x12x10
32@6x6x5
64@6x6x5
64@3x3x2
256
Fully-connected
Vascular reserve
Normal
Decreased
2

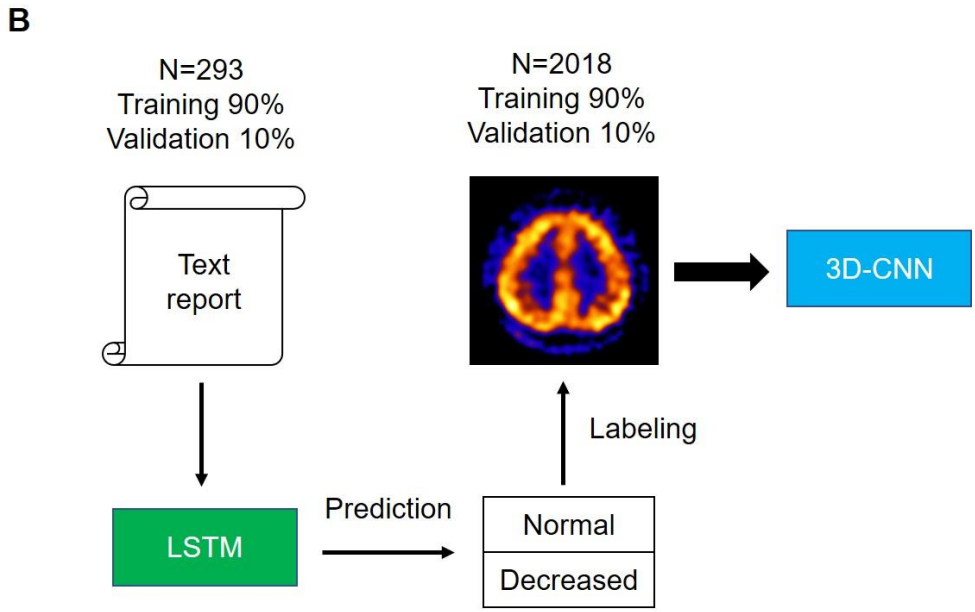


Figure 2. Deep learning-based interpretation of brain perfusion SPECT images

(a) The architecture of 3D CNN framework. Both basal and acetazolamide challenged brain perfusion SPECT images of each patient were used in two different input layers. The images were passed by 3 convolution layers followed by ReLU activation function and max-pooling layer sequentially. The final output of 3D CNN model had two nodes, which correspond to the cerebrovascular reserve state.

(b) The strategy for deep learning-based interpretation of brain perfusion SPECT images. The LSTM model-predicted cerebrovascular reserve of the whole brain was used for the label of each basal/acetazolamide stress brain perfusion SPECT image to train 3D CNN model.

Table 1. Demographics of basal/acetazolamide stress brain perfusion SPECT images used for 3D CNN training

	Total images (n = 2018)
Age at scan date (y)	51.9 ± 17.7 (range 3.0-88.0)
Sex (Male:Female)	952:1066 (47%:53%)

Results

Part I. Performance for text classification of formal reports

The training curve of our designed LSTM model is shown in figure 3. The total accuracy of our LSTM network model for the conclusion classification of text reports regarding given 14 categories of basal perfusion and vascular reserve decrease were evaluated. LSTM model successfully converged and showed 94.1% accuracy to classify text reports into its hemodynamic abnormality (Fig 3a). From the plot of loss, the validation loss was similar but slightly higher than training loss, demonstrating that the model has comparable performance on both train and validation datasets (Fig 3b).

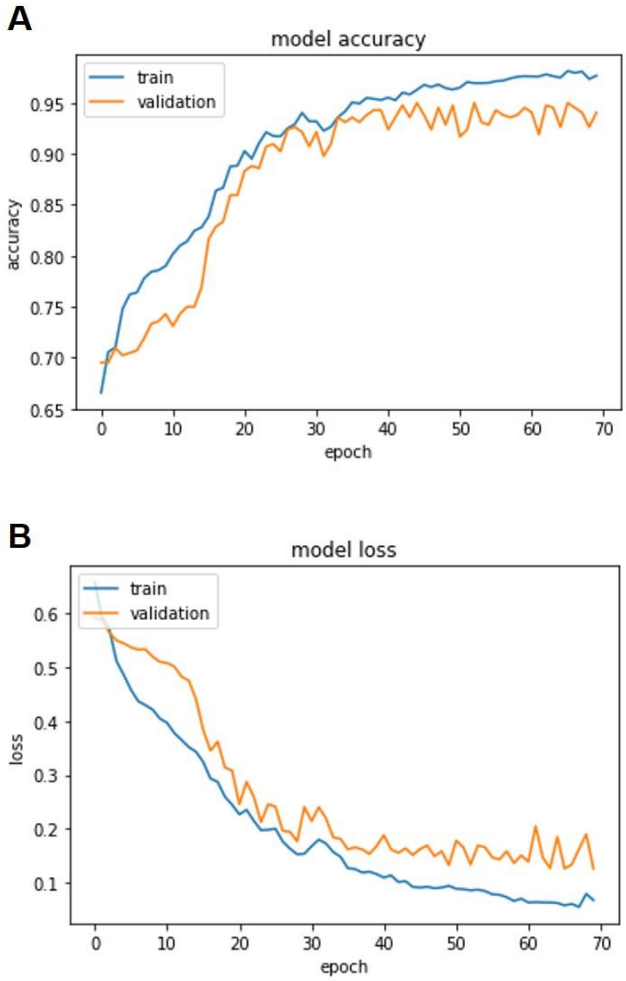


Figure 3. Performance of LSTM network in text classification

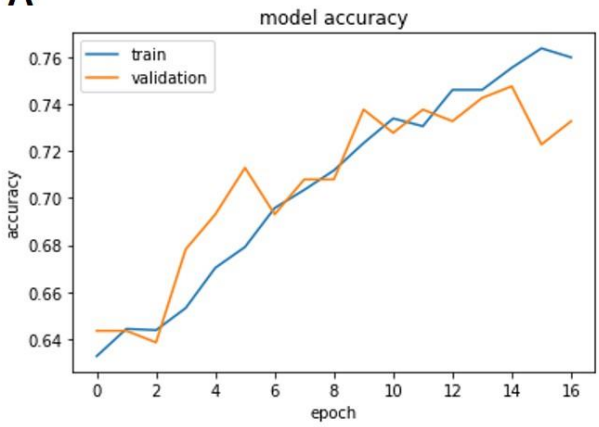
(a) The plot of LSTM network model accuracy on train and validation datasets. (b) The plot of LSTM network model loss on training and validation datasets. (Blue line: train set, Orange line: validation set)

Part II. Performance for image classification of cerebrovascular reserve decrease detection

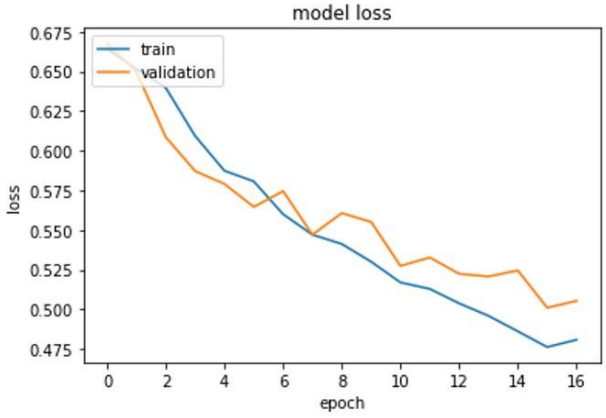
Our trained LSTM model was used to interpret text reports to predict its perfusion abnormality. The LSTM model-predicted cerebrovascular reserve of the whole brain was used for the label of each of basal/acetazolamide stress brain perfusion SPECT image set to train 3D CNN model. The overall strategy for the image interpretation using 3D CNN is summarized in figure 2b.

As shown in the figure 4a, the training set accuracy gradually increased with the number of epochs increases, but the validation set accuracy no longer increased over 74~76% after 17 epochs of training (Fig 4a, 4b). Further training with more epochs increased training set accuracy, but decreased validation accuracy and increased validation loss. To evaluate binary classification performance, receiver operating characteristic (ROC) curve is drawn for 3D CNN model (Figure 4c). The area under the curve (AUC) was 0.81.

A


model accuracy
train
validation
accuracy
epoch
0.76
0.74
0.72
0.70
0.68
0.66
0.64
0
2
4
6
8
10
12
14
16

B


model loss
train
validation
loss
epoch
0.675
0.650
0.625
0.600
0.575
0.550
0.525
0.500
0.475
0
2
4
6
8
10
12
14
16

C

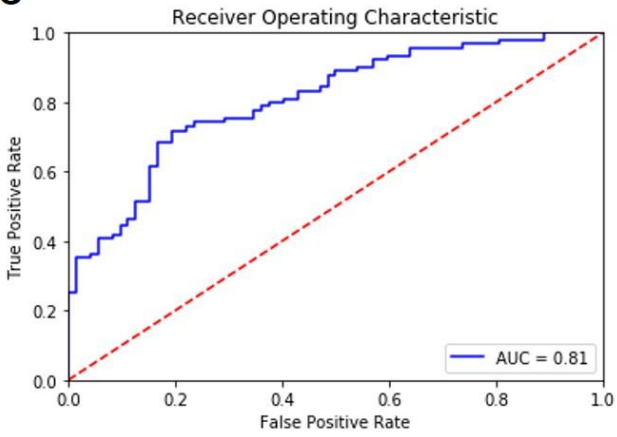

Receiver Operating Characteristic
AUC = 0.81
True Positive Rate
False Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

Figure 4. Performance of 3D CNN in image classification

(a) The plot of 3D CNN model accuracy on train and validation datasets. (b) The plot of 3D CNN model loss on training and validation datasets. (Blue line: train set, Orange line: validation set) (c) The ROC curve for 3D CNN model. Area under curve was 0.81.

Discussion

In this study, we showed the performance of our designed deep learning-based basal/acetazolamide stress brain perfusion SPECT interpretation model. LSTM network showed high accuracy in classifying the conclusion of text reports of the image. The 3D CNN model, however, was trainable but did not showed high performance for image classification to detect the cerebrovascular reserve decrease.

Recently, there are many trials in application of deep learning in medical image classification, segmentation and CNN is most widely used (12, 13, 15). CNN can process image data that come in the form of multiple arrays, including 3D volumetric images used in this study. Typical CNN is composed of a series of multiple convolution and pooling layers and the images are labelled to train the networks.

Recurrent neural network (RNN), a type of deep learning, is used to process sequential inputs such as language model (16). LSTM network, a type of RNN with advances, is suitable to learn about the context of the content in the document for the text classification (17, 18). Several studies using LSTM are performed using time-series medical data or medical image reports as potential resources for deep learning (19, 20). In our study, we both employed CNN and LSTM network for images and text reports to make deep learning-based brain perfusion SPECT interpretation model.

Brain perfusion SPECT with and without acetazolamide challenge is generally

considered to be the gold standard to detect a cerebrovascular reserve decrease. Cerebrovascular reserve evaluation is clinically important to predict the prognosis of cerebrovascular diseases and determine whether the patients will benefit from surgical treatment or interventions (21). The visual assessment of cerebrovascular reserve which is the general method in clinical practice, however, is hardly reproducible due to its interobserver variability. Deep learning-based interpretation system of medical image is expected to overcome inconsistency and interobserver variability of visual interpretation. To our knowledge, this is the first study to attempt deep learning-based interpretation of cerebrovascular reserve on basal/acetazolamide stress brain perfusion SPECT images.

The performance of our designed 3D CNN model, however, should be improved for clinical implementation. The image inhomogeneity, variability and noises are suspicious factors for its difficulty in 3D CNN training. Different cerebral hemodynamic between adult and children or the post-operative hyperperfusion also might cause heterogeneity of input images. Previously reported studies with a successful convolutional neural network performance in medical brain image classification or segmentation used preprocessing step such as spatial normalization or skull stripping (22, 23). Additional image preprocessing procedures are recommended to improve performance of our training models to overcome these challenges. In addition, instead of using typical CNN, state-of-art architecture of deep learning can improve performance.

The superior text classification score of our designed LSTM model show that the

LSTM model-predicted text reports have potential to be used as a feeding label of medical images to train neural networks. However, the text report itself has limitation because it is a subjective data based on human visual interpretation which can result incorrect labeling of the images.

In the clinical setting, not only detecting hemodynamic abnormality but also identifying the region and severity are important. Disease location segmentation and disease severity classification deep learning model should be further studied to get clinical significance. Furthermore, to prove the utility of the deep learning-based interpretation model in the clinical setting, the image interpretation performance should be compared with human visual interpretation score. If the deep-learning based interpretation model have superior performance than unskilled doctor, deep learning-based interpretation system is expected to have a promising role to support doctors in clinical practice.

Even though our 3D CNN model did not show the best performance with several limitations, this study has its own clinical meaning for its first attempt to interpret basal/acetazolamide stress brain perfusion SPECT using deep learning-based methods and to show 3D CNN is trainable using LSTM-predicted text report. More improved performance is expected with an additional image pre-processing step and state-of-art network architecture. If 3D CNN model shows upgraded performance in the future study, deep-learning based interpretation of basal/acetazolamide stress brain perfusion SPECT is expected to have a promising role in assistance of human visual interpretation.

Conclusions

In this study, we designed the deep learning-based interpretation model for basal/acetazolamide stress brain perfusion SPECT image. The LSTM network model show high accuracy in classifying text reports and its predicted data were used to label each basal/acetazolamide stress brain perfusion SPECT image set. The 3D CNN model was trainable, but did not show high performance for the detection of cerebrovascular reserve decrease on brain basal/acetazolamide stress perfusion SPECT image. Our results suggest that 3D CNN is a trainable model for the cerebrovascular reserve decrease detection on basal/acetazolamide stress brain perfusion SPECT using text report prediction of LSTM model as ground truth labels. If we can improve performance with an additional preprocessing step and advanced network architecture, the deep learning-based interpretation is expected to have a promising role in assistance of human visual interpretation in the future.

References

1. Ustrell-Roig X, Serena-Leal J. Stroke. Diagnosis and therapeutic management of cerebrovascular disease. Revista espanola de cardiologia. 2007;60(7):753-69.

2. Klein GM, Seland TP. Occlusive cerebrovascular disease in young adults. Canadian journal of neurological sciences. 1984;11(2):302-4.

3. Nagaraja D, Verma A, Taly A, Kumar M, Jayakumar P. Cerebrovascular disease in children. Acta neurologica scandinavica. 1994;90(4):251-5.

4. Holman BL, Devous MD, Sr. Functional brain SPECT: the emergence of a powerful clinical method. Journal of nuclear medicine. 1992;33(10):1888-904.

5. Baird AE, Austin MC, McKay WJ, Donnan GA. Sensitivity and specificity of 99mTc-HMPAO SPECT cerebral perfusion measurements during the first 48 hours for the localization of cerebral infarction. Stroke. 1997;28(5):976-80.

6. Choi Y. Diamox-enhanced Brain SPECT in Cerebrovascular Diseases. Nuclear medicine and molecular imaging. 2007;41(2).

7. Slomka PJ, Radau P, Hurwitz GA, Dey D. Automated three-dimensional quantification of myocardial perfusion and brain SPECT. Computerized medical imaging and graphics. 2001;25(2):153-64.

8. So Y, Lee HY, Kim SK, Lee JS, Wang KC, Cho BK, et al. Prediction of the clinical outcome of pediatric moyamoya disease with postoperative basal/acetazolamide stress brain perfusion SPECT after revascularization surgery. Stroke. 2005;36(7):1485-9.

9. Choi H, Yoo MY, Cheon GJ, Kang KW, Chung J-K, Lee DS. Parametric Cerebrovascular Reserve Images Using Acetazolamide 99m Tc-HMPAO SPECT: A Feasibility Study of Quantitative Assessment. Nuclear medicine and molecular imaging. 2013;47(3):188-95.

10. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436.

11. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems; 2012.

12. Milletari F, Navab N, Ahmadi S-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. 3D Vision (3DV), 2016 Fourth International Conference on; 2016: IEEE.

13. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention; 2015: Springer, Cham.

14. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. 2014.

15. Li Q, Cai W, Wang X, Zhou Y, Feng DD, Chen M. Medical image classification with convolutional neural network. Control Automation Robotics & Vision (ICARCV), 2014 13th International Conference on; 2014: IEEE.

16. Mikolov T, Karafiát M, Burget L, Černocký J, Khudanpur S. Recurrent neural network based language model. Eleventh Annual Conference of the International Speech Communication Association; 2010.

17. Hochreiter S, Schmidhuber J. Long short-term memory. Neural computation.

1997;9(8):1735-80.

18. Sundermeyer M, Schlüter R, Ney H. LSTM neural networks for language modeling. Thirteenth annual conference of the international speech communication association; 2012.

19. Lipton ZC, Kale DC, Elkan C, Wetzel R. Learning to diagnose with LSTM recurrent neural networks. arXiv preprint arXiv:1511.03677. 2015.

20. Wang X, Peng Y, Lu L, Lu Z, Summers RM. Tienet: Text-image embedding network for common thorax disease classification and reporting in chest x-rays. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2018.

21. Yamamoto K, Miyata T, Momose T, Nagayoshi M. Reduced vascular reserve measured by stressed single photon emission computed tomography carries a high risk for stroke in patients with carotid stenosis. International angiology. 2006;25(4):385.

22. Akkus Z, Galimzianova A, Hoogi A, Rubin DL, Erickson BJ. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. Journal of digital imaging. 2017;30(4):449-59.

23. Choi H, Ha S, Im HJ, Paek SH, Lee DS. Refining diagnosis of Parkinson's disease with deep learning-based interpretation of dopamine transporter imaging. NeuroImage : Clinical. 2017;16:586-94.

국 문 초 록

딥러닝을 이용한 기저/아세타졸아미드 부하 뇌혈류 SPECT에서 뇌혈류 예비능 평가

유현지

서울대학교

융합과학기술대학원

분자의학 및 바이오제약학과

뇌혈관 질환은 환자의 사망과 뇌손상을 일으킬 수 있기 때문에 빠르고 정확한 진단이 중요하다. 기저/아세타졸아미드 부하 뇌혈류 단일광자단층촬영(SPECT)은 기저 뇌혈류 및 뇌혈류 예비능 감소를 평가할 수 있는 기능적 영상이다. 실제 임상에서 뇌혈류 SPECT 영상의 판독은 판독의사의 시각적 판단에 의해 이루어지는데, 이는 종종 판독자 간의 해석의 차이 또는 불일치를 일으킨다. 본 연구에서는 LSTM 신경망과 삼차

원 컨볼루션 신경망(3D CNN)을 이용하여 기저/아세타졸아미드 부하 뇌혈류 SPECT의 판독문 및 영상을 딥러닝을 적용하여 해석해보고자 하였다. 본 연구에서 설계한 LSTM 신경망은 판독문의 기저 뇌혈류 및 뇌혈류 예비능 감소 여부를 분류하는데 좋은 성적을 보이며 훈련 가능한 결과를 보였고, 학습된 LSTM 신경망을 통하여 예측된 분류 결과는 기저/아세타졸아미드 부하 뇌혈류 SPECT 영상에 뇌혈류 예비능 감소가 있는지 3D CNN이 학습할 수 있도록 하는 훈련데이터(label)로 사용되었다. 본 연구에서 설계한 3D CNN 모델은 기저/아세타졸아미드 부하 뇌혈류 SPECT에서 뇌혈류 예비능 감소 여부 영상 분류 학습 훈련이 가능한 결과를 보였으나 뛰어난 성적은 보이지 못했다. 본 연구는 LSTM 신경망을 이용하여 학습된 뇌혈류 SPECT 판독문 분류 예측 결과를 3D CNN을 이용한 뇌혈류 예비능 감소 영상 분류를 위한 학습데이터로 사용하였을 때 3D CNN의 훈련이 가능하다는 것을 보여준다. 기저/아세타졸아미드 부하 뇌혈류 SPECT 해석 딥러닝 모델의 성능 개선을 위하여 추가적인 영상 전처리 과정과 보다 최신의 신경망 구조의 적용이 앞으로의 연구에서 필요할 것으로 생각된다.

주요어 : 딥러닝; 뇌혈류 SPECT; 뇌혈류 예비능; LSTM 신경망; 3D 컨볼루션 신경망

학 번 : 2016-24900