

Inception V3를 이용한 뇌 실질 MRI 영상 분류의 정확도 평가

김지율, 예수영*

부산가톨릭대학교 보건과학대학 방사선학과

Accuracy Evaluation of Brain Parenchymal MRI Image Classification Using Inception V3

Ji-Yul Kim, Soo-Young Ye*

Dept. of Radiological Science, Catholic University of Pusan

요 약 의료영상으로 생성된 데이터의 양은 전문적인 시각적 분석 한계를 점점 초과하여, 자동화된 의료영상 분석의 필요성이 증가되고 있는 실정이다. 이러한 이유 등으로 인하여 본 논문에서는 정상소견과 종양소견을 보이는 각각의 뇌 실질 MRI 의료영상을 이용하여 Inception V3 딥러닝 모델을 이용한 종양 유무에 따른 분류 및 정확도를 평가하였다. 연구 결과, 딥러닝 모델의 정확도 평가는 학습 데이터 세트의 경우 90%, 검증 데이터 세트의 경우 86%의 정확도를 나타내었다. 손실률 평가에서는 학습 데이터 세트의 경우 0.56, 검증 데이터 세트의 경우 1.28의 손실률을 나타내었다. 향후 연구에서는 딥러닝 모델의 성능 향상 및 평가의 신뢰성 확보를 위하여 공개된 의료영상의 데이터를 충분히 확보하고, 라벨링 분류 작업을 통한 라벨링의 정확도를 개선하여 모델링을 구현해 볼 필요가 있다고 사료된다.

• 주제어 : 딥러닝, 인셉션 V3, 케라스, 파이썬, 뇌 실질 MRI 의료영상

Abstract The amount of data generated from medical images is increasingly exceeding the limits of professional visual analysis, and the need for automated medical image analysis is increasing. For this reason, this study evaluated the classification and accuracy according to the presence or absence of tumor using Inception V3 deep learning model, using MRI medical images showing normal and tumor findings. As a result, the accuracy of the deep learning model was 90% for the training data set and 86% for the validation data set. The loss rate was 0.56 for the training data set and 1.28 for the validation data set. In future studies, it is necessary to secure the data of publicly available medical images to improve the performance of the deep learning model and to ensure the reliability of the evaluation, and to implement modeling by improving the accuracy of labeling through labeling classification.

• Key Words : Deep-learning, Inception V3, Keras, Brain MR image

Received 24 September 2019, Revised 28 September 2019, Accepted 30 September 2019

* Corresponding Author Soo Young Ye, Dept. of Radiological Science, Catholic University, 57, Oryundae-ro, Geumjeong-gu, Busan, Korea. E-mail: syye@cup.ac.kr

I. 서론

인간의 모든 생체 기능을 총괄하고 있는 뇌의 구조 및 역할에 대한 의학적 연구는 이미 상당부분 이어져 있으며, 또한 각종 뇌 질환의 발견 및 치유를 위한 의료기술 및 의료영상은 지속적으로 축적 되어 이미 많은 양의 데이터가 있다[1-2]. 최근의 의료영상은 빠른 시간 내에, 저비용으로 얻을 수 있다는 장점이 있으며 [2], 의료영상으로 생성된 데이터의 양은 전문적인 시각적 분석 한계를 점점 초과하여 자동화된 의료영상 분석의 필요성이 증가되고 있는 실정이다. 이러한 이유 등으로 인하여 인공지능의 도입이 가장 활발한 분야가 의료와 헬스케어 영역으로서 의료영상이나 생체 신호를 기반으로 한 질병의 진단에 인공지능 기술의 개발과 도입이 가속화되고 있으며[3], 최근의 딥러닝(Deep-learning) 기술은 다양한 분야에 적용되어 사람들의 예상을 뛰어넘는 결과를 내고 있다[4]. 진단을 위한 의료영상이나 생체신호의 양이 폭발적으로 늘고 있는데 반해, 장기간의 훈련과 임상 경험이 필요한 의료진에 대한 현재의 육성 체계에서는 이러한 수요를 모두 감당할 수 없으며, 의료진 사이의 수련과 임상 경험에 따른 진단 결과의 불 일관성(Inconsistency)이 높다는 점은, 인공지능의 일관성, 확산성, 그리고 정확성과 대비되어 의료 현장 도입에 대한 필요성을 갈수록 높이고 있다[3]. 특히, 이미지 영역을 분류하기 위한 정확하고 신뢰할 수 있는 분할 방법(이미지 영역 분류)은 이미지에서 정보를 추출하기 위한 핵심 요구 사항으로 [5], 딥러닝을 이용한 연구들이 제안되었다. 사전 학습된 오토인코더를 활용한 3차원 합성신경망(3D convolutional neural network, 3D-CNN)이나 사전 학습이 없는 VGGNet, ResNet 등을 기반으로 하는 높은 복잡도의 3차원 합성신경망이 제안되었다[6]. 자연 영상 기반의 딥러닝에서는 모델의 복잡도가 성능과 비례하지만, 의료 영상 기반의 질병 진단 분야에서는 데이터는 크고, 그 수는 매우 많기 때문에 이러한 물이 적용되지 않을 수도 있다. 그러나 이에 대한 체계적인 분석이나 가이드 없이 복잡한 모델들이 계속 제시되고 있다[7-8]. 본 논문에서는 정상소견과 종양소견을 보이는 각각의 뇌 실질 MRI 의료영상 데이터 세트를 이용하여 Inception V3 딥러닝 모델과 케라스(keras) 라이브러리(Library)를 이용한 종양 유무에 따른 분류 및 정확도를 평가하고자 하며, 향후 유사연구 시 기초자료

로 제시하고자 한다.

II. 재료 및 방법

2.1 연구 데이터

Kaggle에서 딥러닝 연구자들을 위해 무료로 제공하는 Brain MRI 의료영상 데이터 세트(<https://www.kaggle.com/navoneel/brain-mri-images>[9])를 연구에 이용하였다. 데이터 세트(dataset)에는 정상소견의 뇌실질 MRI 의료영상 98개, 종양 소견의 뇌실질 MRI 의료영상 155개, 전체 253개의 jpeg 파일로 제공된다. 전체 253개의 데이터 세트를 훈련(training) 163개 데이터, 검증(validation) 50개 데이터, 테스트(test) 40개 데이터로 분류하여 연구에 사용하였다.

2.2 이미지 전처리(Pre-processing Image)

데이터 세트의 뇌실질 MRI 의료영상 데이터는 Statistical Parametric Mapping(SPM)[8]을 사용해서 공간 정규화(Spatial Normalization)한다. 복셀 단위 의 뇌영상 분석에서 뇌 조직의 구조적인 모양과 배치를 기준으로 각 복셀 별 해부학적 위치를 같게 만드는 과정을 공간 정규화라고 한다. 복셀 단위의 뇌 영상 연구에서 가장 기본이 되는 가정은 비교하고자 하는 복셀 위치가 개인별 뇌 영상들에서 같아야 한다는 점이다. 그리고 뇌 이외의 부분은 진단에 필요 없는 부분이므로 crop하여 제거한 후, 픽셀 사이즈 100×100 크기의 이미지를 만든다. 그 후 각 복셀의 값을 0, 1 범위로 배치 정규화한다(Batch Normalization). 배치 정규화는 데이터를 표준화시키는 가장 좋은 방법으로 전체 데이터에 대한 평균과 분산을 사용하는 것이 아니라, 각 계층에 대해서만 평균과 분산에 근사한 값을 찾으므로 효율적 연산이 가능하며, 모든 계층(Layer)의 평균, 크기, 활성화 함수를 독립적으로 조정할 수 있다는 장점이 있다. 그런 다음 Inception V3 input 계층의 이미지 크기에 해당하는 픽셀 사이즈인 299×299 으로 뇌 실질 MRI 영상을 재조정한다.

그림 1은 뇌 실질 MRI 영상의 픽셀 사이즈를 299×299 로 재조정하고 난 후의 이미지 비율 분포를 나타낸 히스토그램이다.

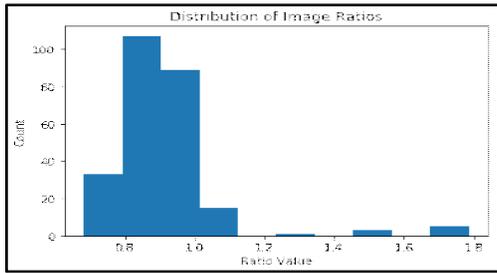


Fig. 1. Distribution of image Ratios

그림 2는 정상소견의 뇌 실질영상의 이미지 전처리 과정을 단계별로 나타낸 그림이다.

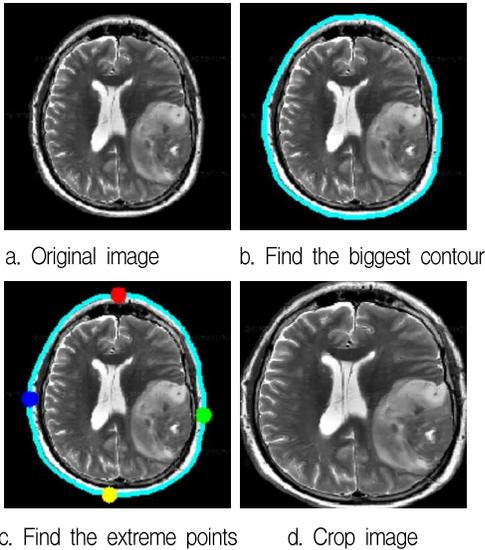


Fig. 2. Imaging pre-processing step

2.3 데이터 확대(Data Augmentation)

데이터 확대는 원본 이미지의 레이블을 변경하지 않고 픽셀을 변화시키는 방법이며, 변형된 데이터를 이용하여 학습을 진행하게 되는데, 인위적으로 변화시킨 이미지는 충분히 학습에 활용될 수 있는 데이터가 된다. 따라서 데이터 확대를 할 경우 적은 데이터 세트로도 대량의 데이터로 증식하여 효과적인 학습 결과를 가져 올 수 있다는 장점이 있다. 그림 3은 데이터 확대기법으로 영상 정규화를 시행한 뇌 실질 MRI 영상이다. 의료영상 데이터는 용량은 크지만, 데이터 수는 한정적이라 수집에 어려움이 있어 이를 분류하는 딥러닝 모델이 항상 좋은 결과를 내는 것에는 어려움이 따른다[8]. 성능평가의 정확성을 높이기 위한 방안으로

모델의 과적합을 해결하기 위해 데이터 개수를 늘리는 정규화(regularization) 기법 중 가장 쉽고 효과가 좋은 데이터 확대 기법[10]을 사용하였으며, 데이터 확대 기법 중 crop을 사용하여 뇌 실질 MRI 의료영상의 데이터를 확보하여 훈련, 검증, 테스트 데이터 세트로 분류하여 실험을 진행하였다.

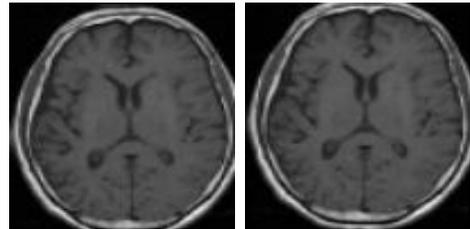


Fig. 3. Augmentation normal brain image

2.4 Inception V3 모델

본 논문에 사용된 모든 CNN 모델의 구조는 그림 4와 같다. 기존의 딥러닝 아키텍처의 경우 계층이 증가하면 연산량이 많아져서 시간이 많이 필요한[11], 반면에 Inception V3의 경우 컨볼루션 레이어(Convolution Layer)의 매개변수의 연산을 줄일 수 있다는 장점이 있다[12]. Inception V3는 그림 4와 같이 22개의 레이어와 21,933,857개의 매개 변수를 포함한다. 이 레이어 중 11개는 1 x 1, 1 x 3, 1 x 5커널의 컨볼루션 레이어이며, 다운 샘플링(Down Sampling)을 위해 컨볼루션 블록이 최대 풀링 레이어(Max Pooling Layer)와 함께 산재되어 있다[10]. 최대 풀링 레이어는 특징 맵의 모든 축의 크기를 1/2로 줄이고 있으며, 이러한 경우 각 블록계층에서 특징 맵의 개수가 2배로 증가한다. 모든 합성곱 계층에서는 배치 정규화기법과 rectified linear unit (ReLU) 계층이 포함되어 있다. 첫 번째 레이어는 Stem layer이며, 모델의 마지막에는 피쳐맵의 크기와 같은 커널 사이즈로 average Pooling을 수행하고, 완전 연결층(Fully connected layer)과 softmax 함수로 1,000개의 클래스에 대한 예측을 수행한다. 파라미터 수를 늘리는 가장 큰 요인 중 하나가 완전 연결층인데 Inception V3의 경우, 이를 적게 사용함으로써 파라미터 수를 절약하여 학습 속도를 향상시키는 장점이 있다.

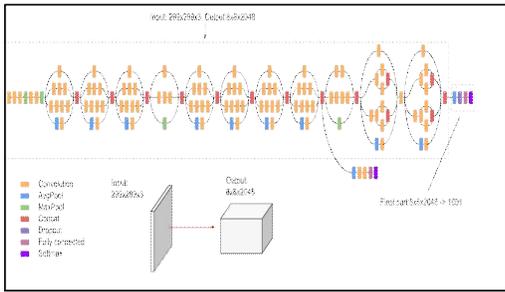


Fig. 4. Inception V3 Model structure

2.5 Keras

파이썬(Python)의 케라스 라이브러리는 연구자가 손쉽게 딥러닝을 구현할 수 있도록 도와주는 상위 레벨의 인터페이스이며, 딥러닝을 쉽게 구현할 수 있다는 장점이 있다[13].

2.6 모델 성능평가

Inception V3를 이용한 모델 성능평가는 뇌 종양 유무에 따른 뇌 실질 MRI 데이터의 실제값에 대한 예측값의 정확도 평가와 epoch에 따른 실제값에 대한 예측값의 손실률 평가를 이용하여 평가하였다.

III. 결과

본 논문에서는 정상조건과 종양조건을 보이는 각각의 뇌 MRI 의료영상 데이터셋을 이용하여 Inception V3 딥러닝 아키텍처와 keras 라이브러리를 이용한 종양 유무에 따른 분류 및 정확도 평가를 하고자 하였으며, 결과는 다음과 같다.

3.1 모델 성능평가 - 정확도 평가

그림 5는 훈련 데이터 세트와 검증 데이터 세트의 종양 유무에 따른 분류 및 정확도 평가를 나타낸 그림이다. 딥러닝 모델의 정확도 평가는 훈련 데이터 세트의 경우 90%, 검증 데이터 세트의 경우 86%의 정확도를 나타내었다.

3.2 모델 성능평가 - 손실률 평가

그림 6은 훈련 데이터 세트와 검증 데이터 세트의 딥러닝 모델링의 손실률을 나타낸 그림이다. 딥러닝 모델링의 손실률 평가는 훈련 데이터 세트의 경우 0.5638, 검증 데이터 세트의 경우 1.2766의 손실률을

나타내었다.

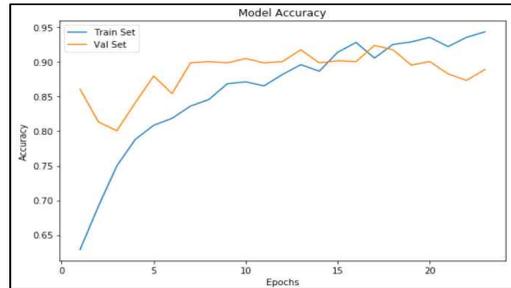


Fig. 5. Modelling Accuracy Evaluation

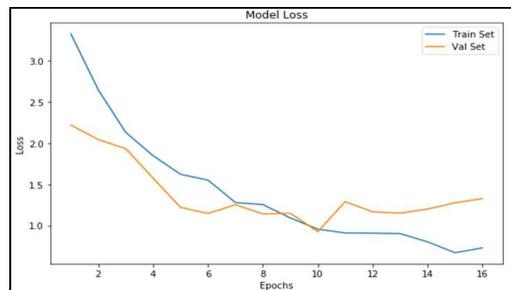


Fig. 6. Modelling Loss Evaluation

3.3 검증 Matrix 실측 및 예측 정확도 평가

그림 7은 검증 데이터 세트의 Matrix 실측 및 예측 정확도 평가를 나타낸 그림이다. 정상조건인 뇌 실질 MRI 데이터의 경우, 전체 18건 중 15건은 정상조건, 3건은 종양조건으로 예측하여 83.3%의 정확도를 확인할 수 있었다. 종양조건인 뇌 실질 MRI 데이터의 경우 전체 32건 중 28건은 종양조건, 4건은 정상조건으로 예측하여 87.5%의 정확도를 확인할 수 있었다.

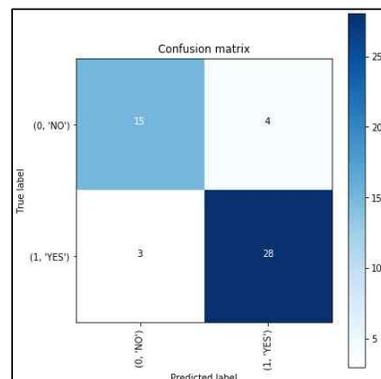


Fig. 7. Validation dataset matrix Accuracy Evaluation

3.4 테스트 Matrix 실측 및 예측 정확도 평가

그림 8은 테스트 데이터 세트의 Matrix 실측 및 예측 정확도 평가를 나타낸 그림이다. 정상소건의 뇌실질 MRI 데이터의 경우 전체 20건 중 18건은 정상소건, 2건은 중앙소건으로 예측하여 90%의 정확도를 확인할 수 있었다. 중앙소건의 뇌실질 MRI 데이터의 경우 전체 20건 중 18건은 중앙소건, 2건은 정상소건으로 예측하여 90%의 정확도를 확인할 수 있었다.

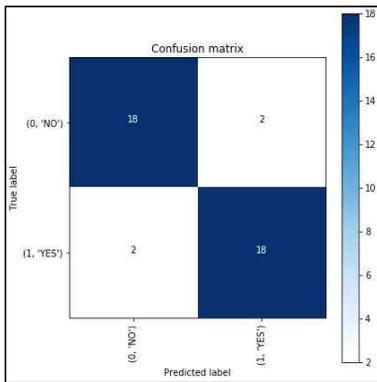


Fig. 8. Test dataset matrix Accuracy Evaluation

IV. 결론 및 고찰

본 논문에서는 Inception V3 아키텍처를 이용하여 뇌실질 MRI 의료영상의 딥러닝 학습을 통한 뇌실질 내 중앙 유무의 정확성을 평가하였다. 딥러닝 모델의 정확도 평가에서는 딥러닝 모델의 정확도 평가는 훈련 데이터 세트의 경우 90%, 검증 데이터 세트의 경우 86%의 정확도를 나타내었다. 딥러닝 모델링의 손실률 평가는 훈련 데이터 세트의 경우 0.5638, 검증 데이터 세트의 경우 1.2766의 손실률을 나타내었다. 이러한 결과는 의료영상이 가지는 한계를 나타낸 결과라고 판단된다. 의료영상은 개인정보보호 등의 이유로 획득에 어려움이 있으며[14], 숙련된 전문가가 아니면, 의료영상의 판독에 어려움 또한 있다. 이러한 이유로 데이터 세트의 충분한 확보에 어려움이 있으며, 라벨링 과정에서의 부정확성이 딥러닝 모델링의 정확도 및 손실률 평가에 영향을 미쳤다고 판단된다. 향후 연구에서는 딥러닝 모델의 성능향상 및 평가의 신뢰성 확보를 위하여 공개된 의료영상의 데이터를 충분히 확보하고, 라벨링 분류 작업을 통한 라벨링의 정확도를 개선하여 모델링을 구현해 볼 필요가 있다고 사료 된다. 본 연

구의 결과는 향후 본 연구와 유사한 기초 연구 시, 연구방법 및 성능평가에 있어 기초연구자료로 제시하고자 한다.

REFERENCES

- [1] Sin hong Kim, "Brain MRI semi-automatic segmentation algorithm for medical image contents," The Korea Contents Society, vol. 5, no. 3, pp. 45-51, Jun, 2005.
- [2] Da him Choi, Seongeun Hong, "Medical Image based Cancer Prognosis Prediction AI Technology : Mainly on lung cancer,"The Journal of The Korean institute of Communication Sciences, vol. 36, no. 4, pp. 10-18, April, 2019.
- [3] Gyuhan Jung, "Current Address and Future Prospects of Artificial Intelligence Technology," Korea Internet Self-governance Organization, vol. 33, pp. 11-15, December, 2018.
- [4] Yeong mo Kim , Jung Woo Lee, "Diagnosis of brain tumor with deep learning,"Proceedings of Symposium of the Korean Institute of communications and Information Sciences, vol 144, No. 3, pp. 1156-1157, July, 2018.
- [5] Rolf A. Heckemann, Joseph V. Hajnal, Paul Aljabar,"Automatic anatomical brain MRI segmentation combining label,"NeuroImage, vol. 33, pp. 115-126, 2006.
- [6] Mingyeong Joen, Hyeonseok Lee1, Sungchan Kim "Empirical analysis of the relation between the complexity and accuracy of CNN models for the diagnosis of Alzheimer disease based on MRI," The Korean Institute of Information Scientists and Engineers, pp. 902-904, July, 2018.
- [7] Alzheimer's Disease International (ADI), "The benefits of early diagnosis and intervention", Alzheimer Disease International. World Alzheimer Report 2011, 2011.
- [8] M. Liu, D. Zhang, E. Adeli-Mosabbeeb, and D.Shen, "Inherent structure based multi-view learning with multitemplate feature representation for Alzheimer's disease diagnosis", IEEE Transactions on Biomedical Engineering, vol. 63, No. 7, pp. 1473-148, 2015.
- [9] <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>, Dataset.

- [10] Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv preprint arXiv, pp. 1409-1556, 2014.
- [11] J.Y. Kim, S.J. Ko "Comparison of DICOM images and various types of images", The Journal of Korea Institute of Signal Processing and Systems, vol. 18, No. 2, pp. 76-83, 2017.
- [12] J M. Havaei et al., "Brain Tumor Segmentation with Deep Neural Networks," Medical Image Analysis, Vol. 35, pp. 18-31, 2017.
- [13] S. Rathore, M. Habes, MA. Iftikhar, "A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. NeuroImage. vol. 155, pp. 530-548, 2017.
- [14] Gyeong u Kim, Jin kuk Kim, "Technical Trends of Medical Image Diagnostic Equipment," The Korean Information Display Society, vol. 19, No. 4, pp. 18-23.

저자 소개

김 지 율 (Ji-Yul Kim)



2015년 3월~현재 :
부산가톨릭대학교 대학원
방사선학 전공 박사과정 중
관심분야 : 디지털 의료영상,
의료영상 딥러닝, 파이썬

예 수 영 (Soo-Young Ye)



1999년 3월 : 부산대학교
의공학협동과정 공학박사
2002년 ~현재 : 부산가톨릭대학교
방사선학과 교수
관심분야 : 인체해부학, 전기전자,
일반화학, 디지털 의료영상처리