

목 차

| | |
|------------------------|----|
| 1. 과제 배경 및 목표 | 1 |
| 1-1. 과제 배경 | 1 |
| 1-2. 과제 목표 | 1 |
| 2. 요구 조건 분석 | 2 |
| 3. 설계 | 3 |
| 3-1. 개발 환경 | 3 |
| 3-2. 데이터 | 3 |
| 3-3. 시스템 구성도 | 9 |
| 4. 개발 일정 및 역할 분담 | 9 |
| 4-1. 개발 일정 | 10 |
| 4-2. 역할 분담 | 11 |
| 5. 참고 자료 | 11 |

1. 과제 배경 및 목표

1-1. 과제 배경

뇌에 악성 종양이 발생할 경우 환자의 생명을 크게 위협할 수 있다. 교모세포종으로 알려진 이 질환은 성인에게서 가장 흔한 형태의 뇌암이자 치명률이 제일 높다. 또한 교모세포종은 최악의 예후를 보이는데, 평균적으로 1년 미만의 생존 기간을 보인다. 그러나 최근 종양에 특정 유전자 서열이 존재할 경우 환자들의 예후가 좋아진다는 점이 발견되었다.

현재 암의 유전자 분석은 조직 샘플을 추출하는 단계가 필요하다. 추출 이후 종양의 유전적 특성을 결정하는데 많은 시간이 소모되고, 선택한 초기 요법의 유형과 결과에 따라 후속 수술이 필요할 수도 있다.

따라서 영상화를 통해 암의 유전을 예측하는 방법이 개발된다면, 환자의 치료 방법을 개선하고 수술 횟수를 최소화할 수 있다. 이를 통해 환자는 수술 전 새롭고 맞춤형 치료를 받을 수 있고, 이는 곧 환자의 관리, 생존 및 전망이 개선될 수 있음을 의미한다.

1-2. 과제 목표

본 졸업과제는 뇌 MRI 영상 데이터와 MGMT 프로모터 메틸화 존재 여부와의 관계를 분석하여 교모세포종의 유전적 아형을 예측하는 것을 목표로 한다.

2. 요구 조건 분석

- 데이터 수집
 - MRI 영상 데이터 수집(dcm 파일)
 - 촬영 방법에 따른 MRI 영상 데이터 분류
- 데이터 전처리
 - 영상 데이터의 이미지 방향, 모양 및 크기가 제각각임
 - 이미지가 같은 공간에 위치할 수 있도록 학습 데이터 처리
- 데이터 학습
 - CNN 방식을 사용한 데이터 학습
 - 데이터 학습 소요 시간 및 정확도 측정
- 데이터 예측
 - 학습된 모델을 바탕으로 하여 입력받은 데이터에 대한 예측
 - MGMT 프로모터 메틸화 여부를 시각적으로 표시

3. 설계

3-1. 개발 환경

- 개발 언어

 - Python

- 사용 라이브러리

 - Pandas(csv 파일 처리)

 - Pydicom(dcm 파일로 구성된 데이터 처리)

 - Matplotlib(dcm 파일의 애니메이션화)

 - Numpy(행렬 및 배열 처리)

 - Keras(딥러닝)

 - Tensorflow(기계 학습)

3-2. 데이터

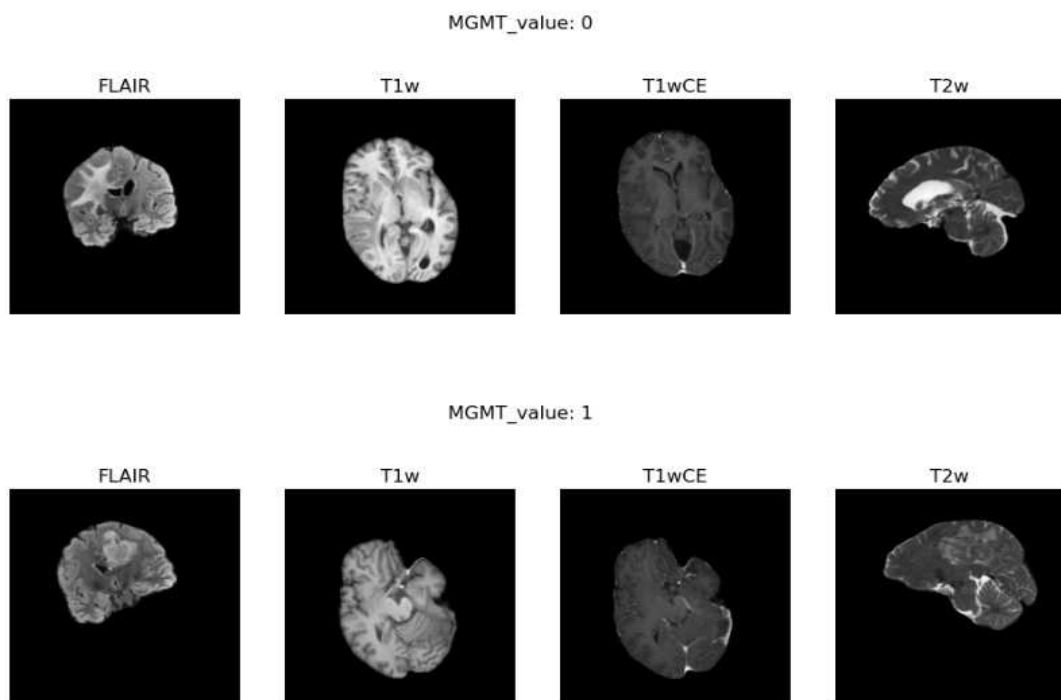
- 데이터 구성

Brain Tumor Segmentation(BraTS) Challenge 에서 제공한 데이터를 바탕으로 학습을 진행할 예정이다. 원본 데이터는 동일한 해상도(1mm^3)를 가지고 두개골이 제거된 형태의 512×512 크기 DICOM 파일 묶음 585종과, 해당 데이터의 메탈화 여부를 알려주는 csv 파일로 구성되어 있다. DICOM 파일 묶음에는 FLAIR, T1 강조(조영제 투입 전), T1 강조(조영제 투입 후), T2 강조 순으로 촬영된 MRI 영상 데이터가 들어있다. T1 강조영상은 TR(Repetition Time)과 TE(Echo Time)를 짧게 하여 얻은 영상으로, 구조적 분석에 용이하다. T2 강조영상은 반대로 TR과 TE를 모두 길게 하여 얻은 영상으로, 병변 분석에 용이하다. FLAIR(FLuid Attenuated Inversion Recovery)는 T2 강조영상의 일종으로, 뇌척수액(CSF)의

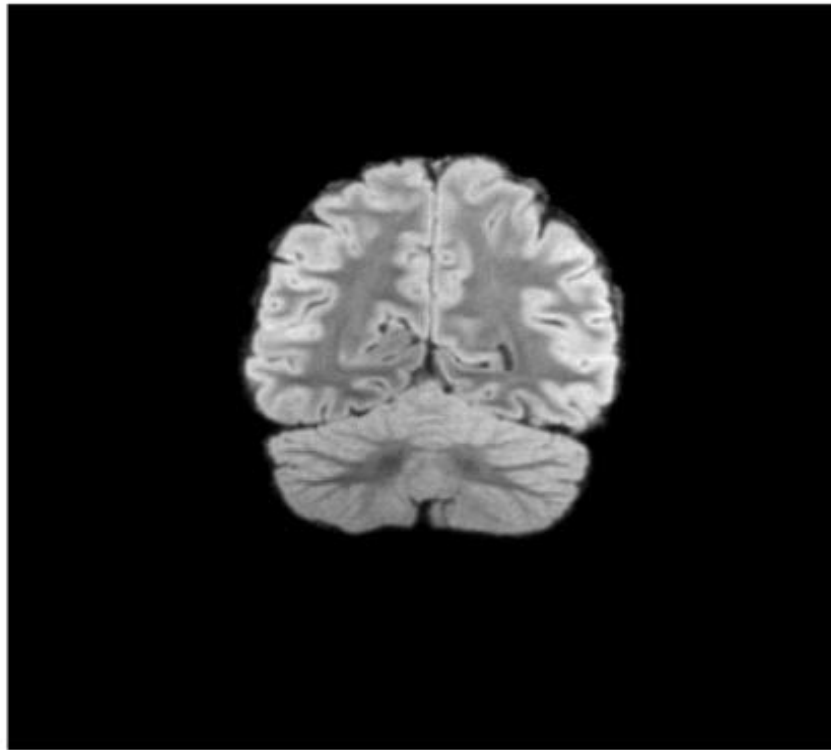
신호를 억제시키는 기법이다. CSF가 검은색으로 나타나 병변을 조금 더 자세히 확인할 수 있다.

- 데이터 확인

학습 데이터들 중 일부와 Matplotlib 라이브러리를 활용하여 <그림 1>, <그림 2>와 같이 데이터를 각각 그림과 애니메이션 방식으로 보여줄 수 있도록 한다.



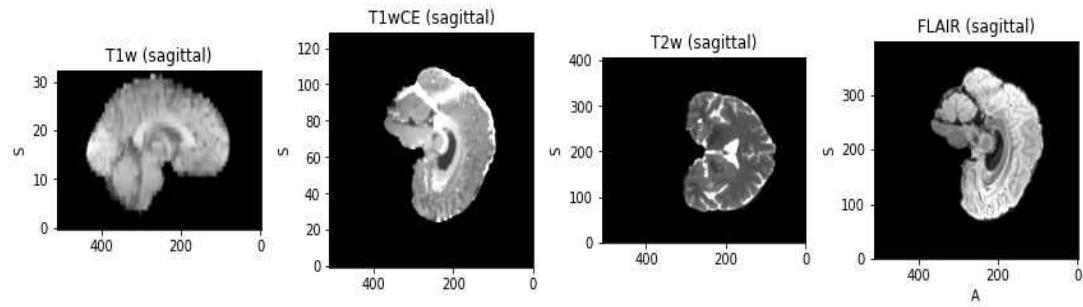
<그림 1> 그림 방식으로 표현한 MRI 데이터



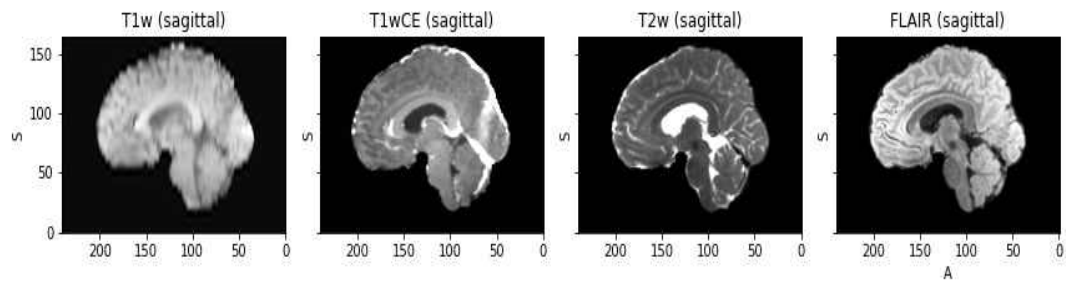
<그림 2> 애니메이션 방식으로 표현한 MRI 데이터

- 데이터 전처리

<그림 3>과 같이 데이터의 방향, 모양, 및 크기가 다른 경우가 확인되어, 본격적인 학습에 들어가기 전 데이터가 <그림 4>와 같이 동일한 공간에 위치할 수 있도록 데이터를 처리한다.



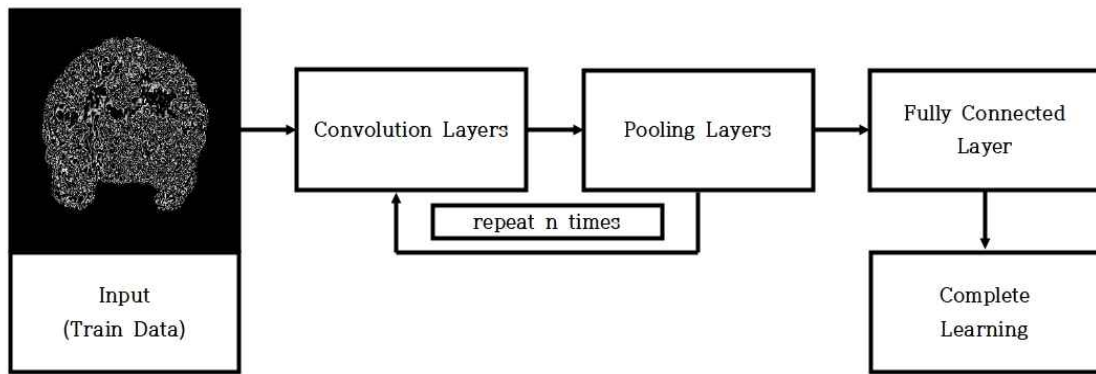
<그림 3> 테스트 데이터(원본)



<그림 4> 테스트 데이터(전처리 후)

- 데이터 학습

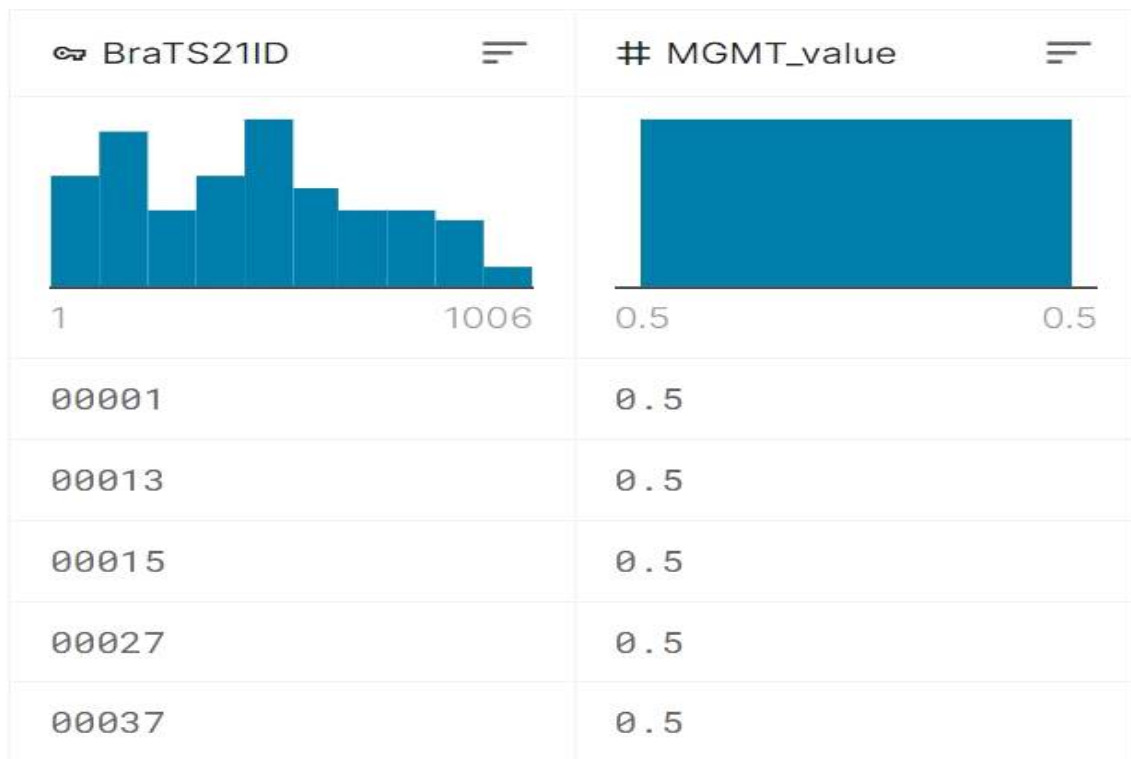
<그림 5>와 같이 CNN 방식을 사용하여 데이터를 학습한다. 코드 테스트와 수정을 여러 번 수행하여 Convolution Layer와 Pooling Layer의 최적 반복 횟수를 찾아보도록 한다.



<그림 5> 학습 모델 구성도

- 데이터 예측

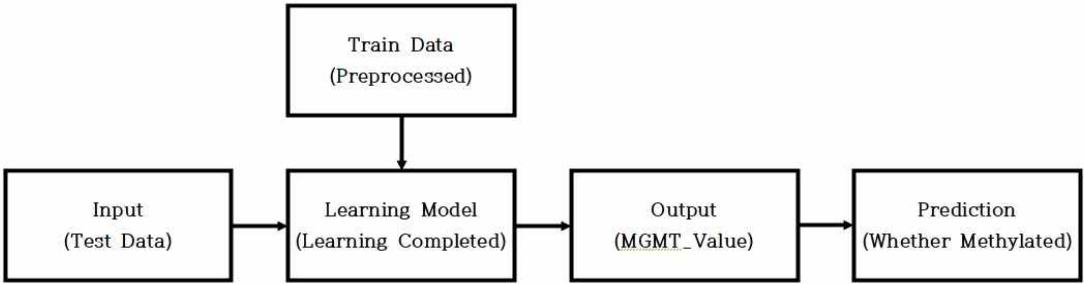
학습이 완료된 후 입력받은 데이터의 예측 결과는 <그림 6>과 같이 표로 나타낼 수 있도록 한다. 표에는 데이터의 번호와 MGMT 메틸화 확률이 포함되어 있다. 만약 입력받은 데이터 종류가 3종류 이하일 경우, <그림 2>에서 표현했던 것과 같이 해당 데이터를 애니메이션 형태로 보여준다. 만약 데이터의 종류가 3종류를 초과할 경우 해당 데이터들을 MGMT 메틸화 확률이 높은 순서대로 정렬 후 상위 3개의 데이터를 애니메이션 형태로 표시한 후, 전체적인 결과를 <그림 6>과 같이 그래프의 형태로 제공할 수 있도록 한다.



<그림 6> 데이터 예측 결과(표 및 그래프)

3-3. 시스템 구성도

전체 시스템의 예상 구성도는 <그림 7>과 같다. 전처리된 학습 데이터를 통해 먼저 학습을
진행한 이후, 메틸화 여부를 알고 싶은 데이터를 입력받는다. 이후, 학습 모델을 통해 해당 데
이터에 대한 메틸화 확률을 계산한 후, 데이터의 메틸화 여부를 결정한다.



<그림 7> 전체 시스템 구성도

4. 개발 일정 및 역할 분담

4-1. 개발 일정

| 5월 | | | 6월 | | | | 7월 | | | | 8월 | | | | | 9월 | | | |
|----------|----|----|----------------|----|----|----|----|----|-----------------|----------------|----|---|----|----|----|----|----|-----------------|----|
| 16 | 23 | 30 | 6 | 13 | 20 | 27 | 4 | 11 | 18 | 25 | 1 | 8 | 15 | 22 | 29 | 5 | 12 | 19 | 26 |
| 관련 지식 습득 | | | | | | | | | | | | | | | | | | | |
| | | | 학습 모델 개발 및 테스트 | | | | | | | | | | | | | | | | |
| | | | | | | | | | 중간 보고서 작성 | | | | | | | | | | |
| | | | | | | | | | | 모델 최적화 및 오류 수정 | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | 최종 테스트 | |
| | | | | | | | | | | | | | | | | | | 최종 보고서 작성 | |

4-2. 역할 분담

| 이름 | 역할 |
|-----|--------------------------------------|
| 이섬재 | 학습 모델 개발 학습 모델 최적화 보고서 작성 및 검토 |
| 한혜린 | 학습 모델 테스트 학습 모델 오류 수정 |
| 공통 | 필요 지식 습득 보고서 작성 및 검토 |

5. 참고 자료

<https://www.med.upenn.edu/cbica/brats2021> (MRI 영상 데이터 출처)

U.Baid, et al., “The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification”, arXiv:2107.02314, 2021.

차이트라 다야난다, 이범식, “CNN을 이용한 뇌 MRI 영상 분할에 관한 성능 비교 분석,” 한국 차세대컴퓨팅학회 논문지, vol.16, no.3, pp. 63-76, 2020.

비젠 카기, “CNN 및 그 유도체를 이용한 뇌 MRI 분할 및 알츠하이머 병 분류,” 학위논문(석사), 조선대학교 일반대학원, 2019.