|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 이름 | 안진영 | 과목명 | 의료 인공지능[7] | 제출일 | 2024.09.12 |
| 학번 | 2293042 | **담당교수** | 정성훈 교수님 | **마감일** | 2024.09.15 |

**의료 인공지능 과제 #1**

**논문 제목: 머신러닝을 이용한 응급실 퇴실 결정 분류 및 주요 인자 추출**

**저자: 황하은**

**- 2022년 2월 고려대학교 석사학위 논문**

1. **서론**

* 응급실의 과밀현상으로 환자의 만족도가 떨어지고 진료의 질에 부정적인 영향을 미쳐왔다. 이를 해결하기 위해 2017년 12월부터 구급차의 분산정책과 진료 인력의 증대, 병상 수 증가 등을 실행했으나 문제는 여전했다.

응급실 과밀은 주로 시술, 환자 입원 및 이송지연으로 발생하고, 퇴원지연은 많은 국가에서 흔히 발생하는 문제이기도 하다. 병원자원을 관리하고 퇴원 정차를 개선하려면 환자 입원을 조기에 예측하는 것이 매우 중요하다. 띠라서 이를 통해서 환자 안전을 확보하고 과밀을 줄일 수 있다.

머신러닝 모델을 활용하여 응급실 퇴실 조치 조기 예측에 대한 다양한 시도가 있었지만, 대부분 연구에서는 단순하게 입원 또는 퇴원이라는 두가지로만 구분하여 이진 분류를 진행하는 한계점을 가지고 있었다.

하지만 이 연구는 지역 의료 환경에 더 잘 맞도록, 국가 응급 의료기관의 데이터를 활용하여 퇴원, 일반 병동, 중환자실 이렇게 세가지 유형으로 분류하여 예측을 향상시키는 것을 목표로 한다.

1. **실험 방법**
   1. **데이터 수집 및 전처리**

* 이 연구는 2016년 1월 1일부터 2018년 12월 31일까지 NEDIS에서 정보를 수집했다. 환자 도착 후 10분 이내에 수집할 수 있는 변수를 중심으로 15세 이하 환자, 결측값이나 이상값이 있는 데이터, 전원 및 사망 환자는 제외하고, 전처리 후 5,818,005개의 기록으로 구성된 데이터 세트를 만들었다.

최종적으로 9개의 연속형 변수와 19개의 범주형 변수를 활용하여, 퇴원 결과를 예측하고, 퇴원 결과를 귀가, 일반 병실 입원, 중환자실 입원 등 세가지 유형으로 분류하여 예측하는 것을 목표로 했다.

<Table 1>은 모든 범주형 독립변수에 따른 귀가, 일반 병실 입원, 중환자실 입원 관측치 개수(%)를 나타낸 표이다. 표에 따르면,

1. 응급 방문율에 요일과 시간대에 따른 비율은 크게 차이가 없다,
2. 여성의 경우 남성보다 귀가의 비율이 높지만 크게 차이는 없다.
3. KTAS 중증도 분류에 따라 퇴실 결정 비율이 상이함을 알 수 있다.
4. 구급차로 도착한 환자의 수가 가장 많았으나 대부분은 퇴원했다.
5. 고급 의료 시설에서 환자를 의뢰했을 때 입원율이 더 높았다.
   1. **분석 방법**

전처리된 데이터를 8:2 비율로 학습 및 평가 세트로 나누고, 5-겹 교차검증을 통해 최적의 매개변수를 선택하여 모델을 구축했다. 사용된 머신러닝 알고리즘은 다항 로지스틱회귀, 의사결정 나무(CART), 랜덤 포레스트, CarBoost, 딥러닝 기반의 TabNet이었다.

1. 다항 로지스틱 회귀는 여러 범주를 예측하는 단순 로지스틱 회귀의 확작된 선형 모델이다.
2. 의사 결정나무는 데이터를 분할하여 패턴을 예측 가능한 규칙들의 조합으로 나타낸다. 본 연구에서는 CART를 사용했다.
3. 랜덤 포레스트는 복원추출을 사용해서 여러 의사결정나무를 결합하여 의사 결정나무의 과적합을 줄인다.
4. CarBoost는 순서형 부스팅 기법을 사용해 범주형 변수를 효과적으로 처리한다.
5. TabNet은 정형 데이터에 적합한 딥러닝 모델로, 어텐션 메커니즘을 통해 의사결정단계에서 변수 선택이 가능하도록 하였다. feature transformer는 입력 데이터를 처리하고 정보를 유지하며, attentive transformer는 이전 단계의 변수 선택을 반영해 마스킹할 변수를 결정한다.
6. **실험 결과**
   1. **실험 평가 지표**

이 연구에서는 성능평가지표로 One vs. Rest AUROC score, accuracy, precision, recall, F1 score을 사용했다.

* 1. **실험 결과**

<Table2>각 모델의 평가 데이터에 대한 성능 결과를 보면, AUROC score 기준으로 모든 모델이 0.77 이상의 성능을 보였다. CatBoost와 TabNet이 모든 평가 지표에서 유사하게 가장 높은 성능을 보였다.

<Figure1> 귀가와 중환자실 입원은 이상적인 결과를 가진다. 일반병실 입원의 경우 귀가와 중환자실 입원으로 다소 혼동하는 모습을 보이지만, 불균형 정도가 심하다는 데이터 특징을 고려하면 어느정도 올바른 결과이다.

<Figure2> CatBoost 모델의 ROC 곡선을 통해 각 범주별 분류성능을 보여준다. (AUROC: ROC 곡선 아래 면적)

1. 중환자실 입원에 대한 AUROC 값은 0.92로, 가장 높은 예측 성능을 보였다.
2. 귀가에 대한 AUROC 값은 0.87로 우수한 성능을 나타낸다.
3. 일반 병실 입원은 0.79로, 상대적으로 낮지만 여전히 신뢰할만한 성능은 보여준다.
   1. **변수 중요도**

CatBoost 모델을 통해 prediction value change를 기준으로 도출된 변수들의 중요도를 분석했다. 최초 KTAS 중증도 경과, 연령, 응급실 내원 시 이동수단이 가장 중요한 변수로 나타났으며, 이것들은 환자의 퇴원 조치결정에 큰 영향을 미쳤다. 특히, 응급실 내원 시간대도 중요한 변수로 선정되었으며, 이는 응급 환자가 많이 발생하는 시간대로 관련이 있었다.

하지만 응급실 내원 년도, 월, 요일 등의 외부 변수는 상대적으로 중요하지 않은 것으로 분석되었다.

TabNet 모델을 통해 각 관측치 별로 글로벌 변수 중요도를 계산한 결과, 30개의 관측치를 샘플링했는데, 변수마다 중요도가 다르게 나타나며 이는 각 환자별로 퇴실결정에 어떤 변수가 중요한지 설명할 수 있는 장점이 있다.

종속변수의 범주에 따른 글로벌 변수 중요도를 보면, 범주 구분 없이 최초 KTAS 중증도 결과가 가장 중요한 변수로 나타났고 산소 포화도 또한 모든 범주에서 높은 중요도를 보였다. 이를 통해 환자의 중증도와 산소 포화도가 퇴실 결정에 중요한 역할을 한다는 것을 확인했다. CatBoost 모델에서도 유사한 상위 변수들이 도출되었음을 언급하며, 변수들이 퇴실 조치 결정에 미치는 영향을 강조했다.

1. **결론**

이 연구에서는 **국가응급진료정보망(NEDIS)** 데이터를 이용해, 환자의 퇴실 조치를 **귀가, 일반 병실 입원, 중환자실 입원**으로 분류했다. 5개의 머신러닝 모델 중 **CatBoost**가 가장 높은 성능(AUROC 0.847)을 보였으며, 특히 중환자실 입원 예측에서 성능이 뛰어났다. (AUROC 0.92)

주요 변수로는 **최초 KTAS 중증도 결과, 응급실 내원 시 이동수단** 등이 도출되었고, 이를 통해 응급실 과밀화 문제를 완화하고 병원 자원 관리에 도움을 줄 수 있다고 결론지었다. 연구의 한계로는 환자의 이전 병력 정보가 포함되지 않은 점을 들었으며, 이를 보완하면 더 정교한 예측 모델을 개발할 수 있을 것으로 예상했다.

**논문 제목: 자기 지도 학습과 어텐션 메커니즘을 활용한 3D 폐 결절 분할 네트워크**

**저자: 정은아**

**- 2024년 2월 중앙대학교 석사학위 논문**

1. **서론**

폐암이 세계적으로 가장 높은 사망률을 기록하고 있지만, 조기 발견 시 치료가 효과적이라는 점을 강조하고 있다. 폐암의 주요 초기 증상 중 하나인 폐 결절을 정확하게 검출하고 분할하는 것이 중요하다고 언급하며 CT 기술을 기반으로 한 결절 검출의 어려움과 방사선 전문의의 수작업 분할 방식의 한계점, 진단의 정확성 문제를 언급한다.

이를 해결하기 위한 방법으로 컴퓨터 보조 진단(CAD) 시스템과 딥러닝 기술을 소개하며, 특히 CNN의 이미지 처리 성능이 의료 영상 분석에 크게 기여하고 있다고 설명합니다. 하지만 기존 지도 학습 기반 모델은 annotated 데이터에 크게 의존하며, 새로운 데이터셋에 적응하는 능력이 부족하다는 문제가 있다. 이에 따라 자기 지도 학습(Self-Supervised Learning, SSL)을 통해 이러한 문제를 해결하고자 하는 연구 방향을 제시하며, 어텐션 메커니즘을 결합하여 모델 성능을 높이는 것을 목표로 한다.

1. **선행연구**
2. Deep Learning Model for Lung Cancer Diagnosis

폐암 조기 발견의 중요성을 강조하며, 폐 결절 검출과 분할을 위한 딥러닝 기반 모델들을 소개한다. 특히, NoduleNet 모델에 대해 설명하는데, 이는 CT 스캔에서 폐 결절을 검출하고 분할하기 위해 개발된 CNN이다. 이 모델은 결절 검출, 거짓 양성 감소, 분할을 동시에 수행하는 엔드-투-엔드 프레임워크이다. 3D CNN으로 구성되며, NCS(결절 검출), DFPR(거짓 양성 감소), SR(결절 분할) 이렇게 3가지 과정으로 이루어져있다. 이는 Multi-task learning을 통해 세가지 작업을 동시에 처리할 수 있지만, shared feature의 비최적화 문제를 방지하기 위해 bounding box를 기준으로 각 작업의 특징을 분리해서 작업을 독립적으로 수행한다.

또한, Faster R-CNN과 같은 모델들이 결절 후보 검출과 분류에 사용되었다. 결절후보검출을 위한 CNN과 거짓 양성 감소를 위한 부스팅 기반 분류기를 결합했다. 기본적인 특징 추출 네트워크로 VGG16을 사용하고, feature map 확대를 위해 deconvolution layer와 두개의 RPN을 추가하여 정확도를 높였다. 특히 2D CT 슬라이스를 입력으로 사용해서 계산 효율성을 증가시켰다. 또한 Dul-Path 모델과 하이브리드 구조가 잇는데, global path와 결절 경계를 강화하는 edge path로 구성된 모델을 설계했다. CNN과 Transformer 기반 하이브리드 구조를 사용하는데, 결절 경계가 모호하거나 흐리게 나타나는 문제를 해결하는데 효과적이다.

1. U-Net

U-Net은 이미지 분할을 위한 신경망 구조로, 픽셀 단위의 정밀한 분석이 필요한 작업에서 효과적이다. U자형 대칭적인 아키텍쳐를 가지고 있고 크게 수축 경로와 확장경로로 나뉜다.

1. 수축경로

U-Net의 왼쪽부분에 해당하는데 Convolution과 풀링 레이어로 이루어져있다. 주로 이지미의 맥락 정보를 학습하는 역할을 한다.

1. 확장경로

U0Net의 오른쪽부분으로, 수축 경로에서 학습한 feature map을 이용해서 업샘플링 과정을 거친다. 여기서는 transpose convolution layer을 사용해서 고해상도의 이미지를 복원하고, 수축 경로에서 얻은 feature map과 결합해서 세부적인 분할 정보를 추출한다.

그리고 U-Net의 핵심적인 구조적 특징 중 하나는 저 두 경로를 연결하는 스킵 연결이다. 수축경로에서 추출한 feature map을 확장 경로의 대응하는 레이어에 직접 연결해서 더 정확한 위치정보를 유지한다.

또한 3D U-Net으로 기존 2D U-Net의 한계를 넘을수도있다. 3D 컨볼루션 레이어를 사용해서 CT이미지의 3차원 정보를 학습하고, 3D MRI나 CT스캔같은 의료 영상에서 더 정밀한 분석이 가능하다. 그리고 V-Net같은 모델은 U-Net을 기반으로 더 깊은 구조에서 안정적인 학습을 위해 Residual Block이나 Dice 손실함수를 사용하기도 한다.

마지막으로 Attention U-Net은 어텐션 매커니즘을 도입해서 더 정밀한 분할을 하기도 한다. 데이터 수집의 어려움을 극복하기 위해 U-Net 구조에 Self-Supervised Learning을 적용한 연구도 진행되고 있다.

1. Self-Supervised Learning

SSL은 레이블 없이 대규모 데이터에서 중요한 특징을 학습하는 방법이다. 핵심은 데이터 자체에서 학습 신호를 추출해서 데이터의 내재적인 구조와 패턴을 학습하는것이다. SSL에서는 학습과정을 Pretext Task와 Downstream Task로 나눌 수 있다.

1. Pretext Task

레이블이 없는 데이터를 통해 모델이 저수준 특징과 고수준 작업별 특징을 학습하게 한다. CNN 모데릉ㄴ 이미지의 가장자리, 텍스쳐 등의 특징을 스스로 학습하게 한다. (ex. 이미지 회전 각도 예측, 이미지 패치의 상대적 위치 예측)

1. Downstream Task

Pretext Task에서 학습한 특징을 바탕으로 실제 문제 해결에 적용된다.

NLP에서는 BERT모델이 SSL을 성공적으로 적용한 사례로, MLM과 NSP 같은 방법을 통해서 텍스트 시퀀스에서 단어의 의미를 학습한다. BERT는 레이블이 필요없는 양방향 in-context learning을 통해서 텍스트의 복잡한 구조를 이해한다.

SSL은 의료영상분석에서도 매우 유용하게 사용되며, Models Genesis라는 모델을 제한해서 3D 의료 영상에서도 SSL을 통한 사전 훈련 가능하다.

1. Attention Mechanism

NLP에서 기계 번역 성능을 향상시키기 위해 도입된 개념이다. 기존의 NMT는 전체 문장을 고정된 길이의 벡터로 압축하는 방식이었지만, 긴 문장에서 정보 손실이 발생하는 문제를 해결하기 위해서 이 매커니즘이 제안되었다. 모델이 특정 이미지 부분에 집중하도록해서 중요한 정보를 효율적으로 학습하는 등 컴퓨터 비전에서도 널리 사용되었다.

그리고 CNN 구조에 어텐션 매커니즘을 결합한 CBAM도 있다. CBAM은 채널 어텐션과 공간 어텐션으로 구성되고, 이를 순차적으로 적용해서 이미지의 중요한 부분을 집중적으로 학습했다. 채널 어텐션이란 CNN의 각 채널이 얼마나 중요한지를 학습하는건데, max pooling과 average pooling을 사용해서 중요한 채널에 더 많은 가중치를 부여하는 방식이다. 그리고 공간 어텐션은 이미지의 어느 위치가 중요한지를 학습해서 특정 위치에 집중할 수 있도록 하는것이다.

또한 Tripler Attention도 있는데, 채널과 공간차원의 상호 의존성을 동시에 고려한 모델이다. 세 개의 병렬 분기를 사용하는데, 채널 어텐션과 공간 어텐션을 분리해서 계산하는 대신에, 채널과 공간 간의 상호작용을 처리하며, 각각의 분기에서 어텐션 맵을 계산한 후에, 이를 결합해서 최종 어텐션 맵을 생성해낸다.

1. **모델**
   1. **NoduleNet**

NoduleNet은 폐 결절을 탐지하고 분할하기 위한 3D CNN이다. 이 모델은 end-to-end 방식으로 설계되었으며, 결절 검출, 거짓 양성 감소, 그리고 분할 작업을 통합하여 CT 스캔을 분석한다. 위에서도 언급했지만, NoduleNet은 세 가지 주요 과정으로 이루어져 있다:

1. 결절 검출(Nodule candidate screening, NCS): CT 스캔에서 잠재적인 결절을 찾아내는 첫 단계로, CNN을 통해 결절 후보 영역을 식별한다.
2. 거짓 양성 감소(Decoupled false positive reduction, DFPR): 시스템의 정밀도를 향상시키는 데 중요한 역할을 하는데, 검출된 결절 중 잘못 검출된 거짓 양성을 제거한다. 이 과정은 분류기를 통해 결절과 비슷한 특성을 가진 혈관 같은 비결절 영역을 필터링한다.
3. 분할(Segmentation refinement, SR): 검출된 결절의 정확한 경계를 정의하는 작업으로 CT 스캔에서 결절의 크기와 모양을 정확하게 분석하여 분할한다.

NoduleNet은 U-Net 아키텍처를 기반으로 한 featurenet 구조로, preBlock, ResBlock3d, BatchNorm3D, ReLU 활성화 함수로 구성되어 있다. RPN은 추출된 특징을 바탕으로 결절이 존재할 가능성이 있는 영역에 대해 bounding boxes를 생성하고, 위치와 크기를 더 정확하게 맞추기 위한 회귀 계수를 계산한다. 또한, 거짓 양성을 줄이기 위해 중간 단계에서 관심 영역인 ROI를 추출하며, RPN과 RCNN 헤드에서 나오는 정보를 결합해 결절의 존재 여부를 평가한다.

마지막으로, RPN의 출력을 기반으로 결절 분할이 진행되며, 원본 CT 이미지에 비례한 크기로 정확한 경계를 정의한다. 기존 NoduleNet은 내부 특징만을 활용했지만, 본 연구에서는 SSL 정보를 추가해서 성능을 향상시키고자 한다.

* 1. **Models Genesis**

3D 의료데이터를 다루는 딥러닝 모델로, SSL을 적용해서 3D 영상 데이터를 학습한다. 기존의 전이 학습 방식을 3D 데이터를 2D 슬라이스로 변환해 처리하면 3D 컨텍스트 정보가 상당부분 소실되는 문제점이 있었으나, Models Genesis를 통해서 이를 해결했다.

1. Pretext Task  
   Models Genesis는 비선형 변환, 로컬 셔플링, 인-페인팅(In-painting), 아웃-페인팅(Out-painting) 등의 변환을 통해 Pretext Task를 수행한다.

이미지는 최대 세가지 변환을 거치지만 인-페인팅과 아웃-페인팅은 동시에 적용될 수 없다.

비선형 변환은 이미지의 밝기에 왜곡을 가하면서 중요정보 잠은 절대적/싱대적 강도 값을 변형시킨다. Modle Genesis는 이러한 변형 이미지를 복원하면서 외형적 특성파악 학습과정을 거친다.

로컬 셔플링은 패치에서 임의부분 샘플링하고, 픽셀 순서 섞어서 변형패치를 생성하는것이다. 인-페인팅과 아웃-페인팅은 모두 context learning에 기여하는데, 인-페인팅은 이미지의 남은 부분을 사용해서 제거된 영역을 복원하고, 아웃-페인팅은 비슷한 방식으로 이미지 외곽 부분을 처리한다.

1. U-Net 기반의 인코더-디코더 구조  
   변형된 이미지를 인코더가 압축하고, 디코더가 이를 재구성해 원본 이미지와 비슷한 이미지를 복원하는 방식으로 학습한다. 이 과정에서 원본 이미지와 재구성된 이미지 사이의 차이를 최소화하는 방향으로 훈련된다.

즉, Models Genesis는 풍부한 3D context를 유지하고, SSL을 통해서 특징들을 완전히 이해하므로 annotated가 없는 데이터로부터 중요한 특징을 학습할 수 있는 모델로 발전하였다.

하지만 이 해당 연구에서는 모델의 범위가 단일 프레임워크로 한전된 제한을 두었고, end-to-end 프레임워크에 적용했을 때의 중요한 특징들을 놓쳤기에, 본 연구에서는 Model Genesis를 end-to-end 프레임워크에 적용할 수 있는 방법을 제시해서 폐 결절 검출, 오진 감소, 분할을 수행할 수 있는 모델을 개발했다.

* 1. **Quartet Attention**

Quartet Attention은 기존 Triplet Attention을 3D 데이터에 적용하여 확장한 개념이다. 이 모듈은 채널 어텐션과 공간 어텐션을 기반으로 하며, 두 종류의 어텐션을 분리해서 처리하는 대신, 채널-공간 차원 간 상호작용을 함께 고려하여 더 정밀한 특징 추출을 가능하게 한다.

Quartet Attention은 C(채널), W(너비), H(높이), D(깊이) 차원 간 상호작용을 학습하기 위해 4차원 텐서 회전을 수행하는 방식으로 동작한다. 각 차원의 상호작용을 학습하는 네 개의 분기 구조를 통해, 공간적 해상도와 채널 의존성 간의 복잡한 관계를 학습한다.

Quartet Attention은 계산 부담을 줄이면서도 다차원 데이터의 복잡성을 효율적으로 처리할 수 있는 메커니즘으로, 의료 영상 분할 작업에서 뛰어난 성능을 발휘합니다.

* 1. **Purposed Model**

Models Genesis에 Quartet Attention Module을 결합하여 새로운 3D 폐 결절 분할 모델을 제안했다

1. SSL을 통한 Pre-train

SSL을 통해서 CT 데이터를 pre-train한다. 이 과정에서 중요한 3D 구조와 텍스쳐를 학습하고, downstream task로 전이학습이 가능하게 된다.

1. Downstream Task

Pre-trained 된 가중치를 바탕으로 NoduleNet의 폐 결절 검출, 거짓 양성 감소, 분할 작업을 수행한다. 그리고 ResBlock마다 Quarter Attention Module을 추가해서 성능을 향상시킨다.

즉, 기존 NoduleNet과는 달리, SSL을 통해서 얻은 정보로 성능을 극대화하면서, 동시에 추가 파라미터를 최소화해서 모델의 복잡성을 줄인것이다.

* 1. **손실함수**

사전 학습 손실 함수로 L1-norm을 사용해서 원본 이미지와 재구성된 이미지 간의 차이를 최소화했다. Downstream Task에서는 RPN 손실함수, RCNN 손실함수, Mask 손실함수가 사용되었다. RPN 손실함수는 Faster R-CNN의 손실 함수와 동일하며, 경계 상자 회귀 계수를 포함한다. RCNN 손실함수는 RPN 손실함수와 동일한 방식으로 계산되고, Mask 손실함수는 Semantic Segmentation작업에서는 soft-dice loss를 사용해서 예측된 분할과 실제 레이블 간의 유사도를 계산하여 모델을 학습시킨다.

1. **실험**
   1. **데이터 소개**

여기서는 **LUNA16** 데이터셋을 사용하여 실험을 진행하였다. LUNA16은 **LIDC-IDRI** 데이터셋에서 추출된 일부로, 1018명의 환자 CT 스캔 데이터를 포함하고 있으며, 네 명의 방사선 전문의가 주석을 달아서 제공한다. LUNA16 데이터셋은 **얇은 슬라이스**의 CT 스캔을 포함하고 있어, 결절 관리에 적합하다. 하지만, 슬라이스 간격이 일관되지 않거나, 두께가 3mm 이상인 슬라이스가 포함된 경우, 혹은 슬라이스가 빠진 경우는 제외된다. 즉, LUNA16은 **888개의 CT 스캔**과 **1186개의 양성 결절**로 구성되어 있다.

연구에서는 SSL을 위한 **사전 학습**과 **Downstream Task를** 위해 이 데이터셋을 사용했다. 레이블이 있는 데이터셋이긴하지만, SSL을 통해 데이터의 특성을 효과적으로 학습할 수 있음을 기대했기 때문이다. SSL 단계에서 CT 이미지는 **64 × 64 × 32** 크기로 추출했고, **Downstream Tas**에서는 **128 × 128 × 128** 크기로 추출했다. 또한, SSL이 주석을 필요로 하지 않는다는 점을 활용해서, **검증 및 테스트 절차에서의 label 노출** 위험을 최소화하면서 데이터셋의 모든 CT 이미지를 사용하였다.

* 1. **평가지표**

**LUNA16 대회**에서 사용된 CPM(Competence Performance Metric)을 기반으로 한 **FROC(Free Response Receiver Operating Characteristic)** 분석을 평가 지표로 사용하였다. 전통적인 ROC 분석과 달리, **FROC 분석**은 병변 감지 시스템의 성능을 평가하는데, True Positive Rate와 False Positive Rate을 동시에 고려한다. 이 분석 방법은 하나의 영상에서 여러 병변을 얼마나 정확하게 감지하는지를 평가하는 데 유용하다.

여기서 Confusion Matrix은 시스템의 성능을 이해하는 데 중요한 역할을 하며, 참 양성(True Positive), 거짓 양성(False Positive), 참 음성(True Negative), 거짓 음성(False Negative)으로 구성된다. 민감도(Sensitivity)는 참 양성 비율을 의미하며, 모델이 실제 결절을 얼마나 잘 감지하는지 평가하는 척도이다. 폐암과 같은 질병에서 높은 민감도는 매우 중요하다. 민감도가 높을수록, 실제 결절을 놓치지 않고 감지할 수 있어, 의료 전문가가 후속 조치를 취할 기회가 많아지기 때문이다.

CPM은 FROC 분석에서 파생된 평가 지표로, **FPs/scan** 값에 해당하는 민감도의 평균으로 계산된다. 본 연구에서는 **FPs/scan** 값이 **0.125, 0.25, 0.5, 1, 2, 4, 8**일 때 각각의 민감도를 평균내어 CPM을 산출하였다. 예를 들어, FPs/scan 값이 8일 때 민감도가 0.964라는 것은 스캔당 평균 8개의 거짓 양성을 허용하는 조건에서 시스템의 민감도가 96.4%임을 의미한다.

* 1. **실험 및 결과**

이 연구에서는 **사전 학습**과 **다운스트림 작업**을 위한 두 가지 실험을 수행하였다.

사전 학습을 위한 초기 학습률은 **1**로 설정되었으며, **Momentum이 0.9**인 SGD을 사용해 훈련을 진행하였다. 다운스트림 작업을 위한 모델은 초기 학습률이 **0.01**로 설정되었고, 동일하게 **Momentum 0.9**의 SGD를 사용하였다. **RCNN**은 65번째 epoch에서 시작되었으며, **최종 분할 작업**은 80번째 에포크에서 시작되었다.

**NoduleNet**과 변형된 Models Genesis(MG)로 사전 학습한 모델 간 성능을 비교한 결과, MG 모델은 사전 학습된 가중치를 사용하여 NoduleNet에 비해 더 나은 성능을 보였으며, 특히 **거짓 양성 감소(DFPR)**와 **결절 분할(SR)**에서 더 나은 성능을 보였다. **FPs/scan 값**을 0.125에서 8까지 다양하게 설정한 조건에서, 변형된 MG 모델은 NoduleNet보다 더 높은 민감도를 기록하였다. 예를 들어, FPs/scan 값이 8일 때, MG 모델은 96.4%의 민감도를 보여, NoduleNet보다 더 많은 결절을 정확히 탐지하였다.

결과적으로, SSL을 통해 **사전 학습된 가중치**를 사용하면, 레이블이 부족한 환경에서도 **정확한 결절 분할**이 가능하다는 점을 확인하였다. 특히 MG 모델은 레이블이 없는 데이터를 활용하면서도 높은 정확도를 유지할 수 있어서 폐 결절 검출 및 분할 작업에서 매우 유용한 것으로 나타났다.

1. **결론 및 향후 연구**
   1. **결론**

본 연구에서는 SSL과 Attention 매커니즘을 결합해서 label이 지정된 데이터가 제한된 환경에서 효율성과 정확성을 개선하는 새로운 3D 폐 결정 분할 네트워크를 제안했다. 특히, 제안된 모델은 사전 학습 단계에서 SSL 방식을 활용하여 레이블이 부족한 상황에서도 우수한 성능을 발휘할 수 있도록 했다.

실험 결과, SSL 방식의 사전 학습과 Attention 메커니즘을 도입한 모델은 모든 FPs/scan 범위에서 균형 잡힌 성능 향상을 보여주었으며, 특히 낮은 FPs/scan 수준에서 높은 민감도를 달성했다. 이 결과는 제안된 모델이 기존 모델들과 비교하여 효율적인 폐 결절 분할을 수행할 수 있음을 입증했다. 또한, 모델의 평균 CPM값이 0.848로, 이는 현재 사용 중인 대부분의 딥러닝 모델보다 우수한 성능을 기록했다​.

* 1. **향후 연구**

향후 연구에서는 label이 없는 대규모 데이터 셋을 활용하여 사전 학습을 수행하고, 이를 바탕으로 다양한 downstream tasks에서의 성능을 평가하여 데이터 활용 능력을 비교할 계획이다. 또한, 제안된 모델의 범용성을 높이기 위해 MRI나 X-ray와 같은 다른 의료 영상 모달리티(Modality)와 뇌, 간, 심장 등 다양한 신체 기관의 데이터에 적용할 예정이다. 이러한 접근은 제안된 방법론이 의료 영상 분석의 다양한 영역으로 확장될 수 있는 가능성을 탐색하고, 효과적인 진단 도구로서의 잠재력을 증명하는 데 중점을 둘 것이다​.