

# 다중 모드 데이터를 사용한 폐암 생존분석 검토

최철웅\*, 김현지\*\*, 심은석\*\*, 임아연\*\*, 이윤준\*\*, 정선주\*\*, 김경백\*

\*전남대학교 인공지능융합학과

\*\*전남대학교 전자컴퓨터공학부

sentilemon02@gmail.com, khyeonj1025@gmail.com, ses8323@gmail.com,

gudwodjssl@gmail.com, dbswns97@gmail.com, ju990317@gmail.com,

kyungbaekkim@jnu.ac.kr

## Review of Lung Cancer Survival Analysis with Multimodal Data

Chul-woong Choi\*, Hyeon-Ji Kim\*\*, Eun-Seok Shim\*\*, A-yeon Im\*\*,

Yun-Jun Lee\*\*, Seon-Ju Jeong\*\*, Kyung-baek Kim\*,

\*Dept. of Artificial Intelligence Convergence, Chonnam National University

\*\*Dept. of Electronics and Computer Engineering, Chonnam National University

### 요 약

폐암 환자의 생존율을 예측할 때 미국암연합회(AJCC)의 TNM병기 분류체계에 의해 진단되는 최종병기를 많이 사용한다. 최종병기는 폐암환자의 임상데이터 중 하나로 종양의 위치, 크기, 전이정도를 고려하여 환자의 폐암 상태를 판별하는 정보이다. 최종병기는 개략적인 환자의 상황을 설명하는데 효과적이지만, 보다 구체적인 생존분석을 위해서는 임상데이터 뿐만 아니라 PET/CT와 같은 영상 데이터를 함께 분석해야 한다. 이 논문에서는 데이터 과학적 접근을 통해 폐암환자의 임상데이터, CT영상과 PET영상 등 다양한 종류의 데이터를 함께 활용하는 생존분석기법을 검토한다. 실험을 통해 다중 모드 데이터를 활용하는 생존분석을 위해 비선형모델 개발과 Feature임베딩 기법 고도화가 필요함을 확인하였다.

### 1. 서론

대한민국 암 발생률을 OECD 평균보다 낮지만 총 사망자의 26.5%가 암 환자이며, 그 중 폐암 환자가 가장 많은 비중을 차지한다[1]. 폐암은 1기에서 4기로 진행되는 속도가 1년이 안될 정도로 빠르게 확산되며 우리나라뿐만 아니라 전 세계적으로 사망률 1위다[2].

폐암환자의 생존율을 예측할 때 임상데이터를 바탕으로 5년, 10년과 같이 큰 범주로 구분하고 있다. 임상데이터의 Feature만으로는 종양의 위치, 크기, 전이정도 등 환자의 다양한 정보를 알 수 없다. 의료용 디지털영상 표준인 DICOM 영상을 임상데이터와 함께 사용해야만 정확한 생존분석이 가능하다[3]. DICOM영상은 의료용 디지털 영상 및 통신표준으로 CT와 PET 등이 있다. DICOM영상은 성인기준으로 환자 1명당 약 300장 정도의 이미지를 갖기 때문에 효과적으로 중요한 Feature를 추출하는 방법 등이 필요하다.

일반적으로 가장 많이 사용하는 생존분석 모델은

통계적 모델인 Cox 비례위험 모형이 있다. 현재는 인공지능 딥러닝 모델의 발달로 생존분석에 Neural Network를 활용한 DeepSurv 모델과 같은 다양한 연구가 진행되고 있다[4].

이 논문에서는 폐암환자의 다중 모드 데이터를 사용하여 Feature를 추출하고 생존분석을 수행하는 기법과 발전 가능성에 대해 전반적으로 검토한다.

2장 관련연구에서는 Cox비례위험모델, DeepSurv와 Image Feature Extraction에 대해 알아보고 3장에서는 멀티모달데이터 기반 폐암 생존분석 기법에 대해 검토한다. 4장에서는 실험을 통해 Cox비례위험 모형과 DeepSurv 생존분석 기법 성능을 비교하고, 5장에서는 결론에 대해 기술한다.

### 2. 관련 연구

#### 2.1 Cox 비례위험 모델

생존분석이란 정의된 시작지점으로부터 어떤 이벤트가 발생하기까지의 시간을 분석하고 예측한다. 이벤트는 기계고장, 사망, 질병악화 등이 있다. 생존분석에는 특정변수의 relative hazard를 계산할 수

있는 통계적 회귀모델인 Cox 비례위험 모형(Cox PH, Cox Proportional Hazard Model)이 많이 사용된다[5].

Cox 비례위험 모델은 존재한다는 가정과 비례위험의 가정을 한 다음 회귀분석을 통해 중도절단(censored)된 자료를 처리한다. 비례위험(PH, Proportional Hazard)은 시간에 관계없이 어떠한 변수의 위험비(Hazard Ratio)는 항상 일정하다는 기본가정을 기반으로 하며, 식 (1)과 같은 생존함수를 사용한다.

$$h(t | X_i) = h_0(t) \exp(X_i \cdot \beta) \quad (1)$$

수식(1)의  $h_0$ 는 위험함수,  $X_i$ 는 대상  $i$ 의 변수, 시간  $t$ ,  $\beta$ 는 모델이 학습해야 할 계수를 의미하며, 특정시간에 이벤트 발생 확률을 계산할 수 있다[6].

Cox 비례위험 모델은 의료분야 뿐만 아니라 중소기업의 생존요인분석[7], 교통법규 위반 억제 효과 분석[8] 등 다양한 분야에서 사용되고 있다

### 2.2 DeepSurv

DeepSurv는 Faraggi-Simon 네트워크와 유사한 다층 퍼셉트론으로 Cox 비례위험 모델의 딥러닝 일반화를 구현한 것으로, 비선형적 비례위험 모델이다. DeepSurv는 공변량의 사전 선택이 필요하지 않지만 적용적으로 학습하므로 기존 Cox 회귀에 비해 이점이 있다.

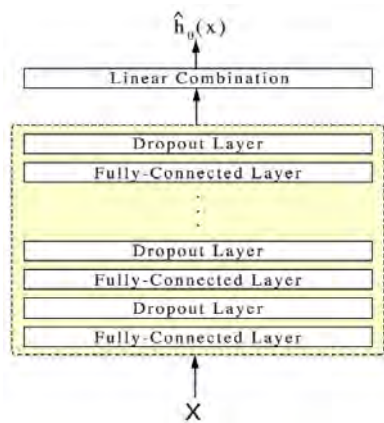


그림 1. DeepSurv 모델 구조

DeepSurv는 deep architecture(즉, 하나 이상의 hidden layer)를 허용하고 weight decay regulation, ReLU(Rectified Linear Units), Batch Normalization, dropout, 경사하강법 등과 같은 기술을 사용한다. 심층 네트워크의 최상위 숨겨진 계층의 공변량이 Cox 비례위험 모델에 대한 입력으로 사용되도록 네트워크에 숨겨진 계층을 추가하여 네트워크를 개선한다.

네트워크의 출력은 네트워크 가중치  $\theta$ 로 매개 변수화 된 위험 함수  $h^\theta(x)$ 를 추정하는 단일노드이다.

DeepSurv는 네트워크 출력노드를 조절하여 비선형적 모델 분포의 간격을 조절할 수 있으며, 단일 데이터 입력에 대한 개별 비선형 분포를 예측할 수 있다. DeepSurv는 다양한 생존분석 애플리케이션에서 사용할 수 있다. 예시로 일련의 환자 관찰에 대한 치료 권장 사항을 제공하는 하나의 의료 응용 프로그램인 recommended\_treatment[4]가 있다.

### 2.3 Image Feature Extraction

DICOM(Digital Imaging and COmmunications in Medicine)은 의학 영상 정보 및 자료의 취급, 저장, 표시 및 전송을 위한 국제 표준이다. 다양한 영상 진단장치(X-ray, CT, PET, MRI)를 통하여 획득된 영상정보를 디지털 상태로 저장하고 전달하려면, 영상들을 DICOM 포맷으로 저장하는 것이 필수적이다.

현재 사용되는 CT(Computed Tomography) 자료는 일반적으로 가로 512픽셀과 세로 512픽셀의 해상도를 가지고 있다.[9] PET(Positron Emission Tomography)은 기능적 영상진단 기술 중 하나로, 포도당과 비슷한 물질인 FDG(Fluorodeoxyglucose)를 신체에 주사한 뒤, FDG가 신체에 흡수되면 양전자 방출 핵종이 붕괴하여 분자 구조가 변하기 전까지 남아 있게 되는데, 이를 핵의학 영상 기기들을 이용해 촬영하면, 대사가 높은 부분을 찾을 수 있다. 이를 CT자료에 매칭하여 해부학적으로 어느 부위에서 암의 전이가 얼마나 일어났는지 알 수 있다.

CT와 PET은 2차원 데이터이다. 하지만, 인체를 모두 표현하기 위해서는 3차원 데이터가 필요하다. 따라서, 한 사람의 전신 CT 또는 PET 정보는 2차원 이미지 200~300 장 정도가 필요하다. 2차원 이미지의 해상도를 512x512픽셀이라고 한다면, 한 사람의 전신 3차원 데이터의 해상도는 512x512x250 ≈ 65백만 픽셀이 된다. CT와 PET을 동시에 고려한다면 한 사람에 대한 정보의 수는 약 1.3억 픽셀이 된다.

이러한 대용량의 이미지 정보를 이용하여 생존분석을 수행하기 위해서는, 정보의 특징을 추출하여 활용할 수 있다. 최근, 이미지 학습을 위한 딥러닝 모델로 CNN(Convolutional Neural Network)이 연구되고 있으며, VGGNet이나 ResNet과 같이 네트워크 구조를 강화하여 특징 추출 성능을 향상시키는 합성곱 네트워크들이 개발되고 있다. 이 연구에서는

최근 각광을 받고 있는 ResNet을 활용하여 영상의 특징을 추출하였다.

VGGNet과 같은 Forward 네트워크로 구성된 합성곱네트워크는 심층학습의 망이 깊어지면, 기울기 소멸(vanishing gradient) 또는 포화(exploding gradient)와 같은 문제 때문에 학습의 성능이 저하되는 문제(degradation problem)가 발생한다. 이러한 문제의 해결을 위해 ResNet은 잔여학습(Residual Learning)개념을 사용한다.[10]

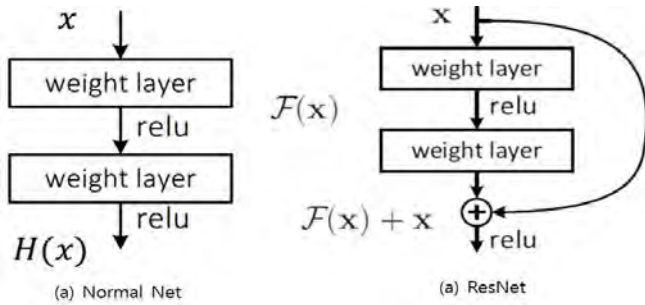


그림 2. Residual Learning 개념 비교

그림 2(a)와 같이 일반적인 학습망에서는  $H(x)$ 가 최적화되도록 학습을 수행한다. 하지만, 그림 2(b)와 같이 잔여학습(residual learning)은 개념을 달리하여  $H(x) - x$ 를 최적화하도록 학습한다. 여기서,  $F(x)$ 는 식 (2)로 정의한다.

$$F(x) = H(x) - x \quad (2)$$

이때,  $H(x) - x$ 의 최적화는  $F(x)$ 가 0으로 수렴하는 것이다. 이는 학습할 방향이 미리 결정되어 있는 사전 조정(pre-conditioning)의 역할을 하게 된다. 즉,  $F(x)$ 가 0이 되는 방향으로 학습하게 되면 입력의 작은 움직임을 쉽게 검출할 수 있게 되며, 식 (2)와 같이 출력과 입력의 차로 학습하기 때문에 잔여학습이라고 부른다.

또한, 입력과 같은  $x$ 가 출력에 연결되기 때문에 파라미터의 수에 영향이 없고, 덧셈이 늘어나는 것을 제외하면 shortcut 연결을 통한 연산량 증가는 없다. 또한, layer를 건너뛰면서 입출력이 shortcut으로 연결되어 forward/backward path가 단순해지는 효과를 얻을 수 있다.

이 연구에서는 CT와 PET과 같은 3차원 인체 영상데이터를 다루게 되어 3D 이미지를 입력으로 받아 특징을 추출할 수 있는 ResNet3D모델을 이용하여 의료영상 이미지 특징을 추출 하였다.

### 3. 다중 모드 데이터 기반 폐암 생존분석 기법

폐암환자에 대한 임상데이터와 영상데이터를 함

께 사용하는 다중 모드(Multi-modal) 데이터 기반 생존분석 기법은 그림 3과 같은 절차로 수행될 수 있다.

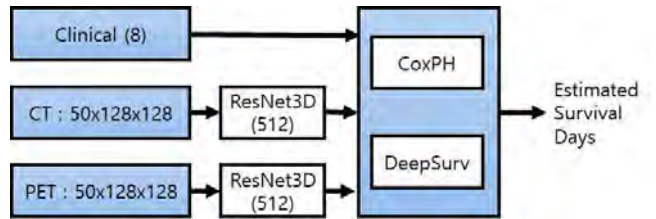


그림 3. 다중 모드 데이터 기반 생존분석 모델 개요

표 1. 데이터셋 설명

Clinical(10)	성별, 나이, 폐암 최종병기, 암 세포의 크기(T), 림프절 전이 정도(N), 타 장기 전이 여부(M), 흡연 여부, 흡연량, 사건(사망) 발생 여부, 생존 시간
CT(512)	ResNet3D를 사용하여 CT이미지에서 추출된 512개 특징
PET(512)	ResNet3D를 사용하여 PET이미지에서 추출된 512개 특징

폐암환자의 임상데이터는 표1과 같이 10개로 구성된다. 이 연구에서는 앞에서 열거한 정보 중 생존 시간과 사건발생여부를 제외한 8가지 정보를 생존분석을 위한 입력 변수로 활용하였다.

영상데이터의 CT와 PET은 표1과 같으며, CT와 PET은 사람의 키에 따라 영상이미지 개수가 달라지는 점을 고려하여, 한사람의 CT와 PET개수는 50개로 표준화 한다. 또한, ResNet3D의 학습 연산을 현실적인 수준으로 맞추기 위해 각 CT와 PET 입력 영상의 해상도를 128x128 픽셀로 조정하였다. 이렇게 조정된 CT와 PET영상을 출력이 512개인 ResNet3D를 통해 학습하고 그 출력을 생존분석을 위한 입력 변수로 활용하였다.

### 4. 성능 검증 및 평가

다중 모드 데이터기반 생존분석 모델의 성능을 평가하기 위해, 폐암환자 184명의 데이터를 활용하였다. 임상데이터 및 영상데이터는 3장에서 설명한 9개의 임상데이터와 CT/PET 영상데이터를 사용하였다. CT/PET 영상에서 추출된 1024 특징 값은 연산량을 줄이기 위해 -1.0~1.0사의 값으로 정규화 하여 사용하였다.

생존분석 모델의 성능평가를 위해 사망여부를 고려한 생존시간 예측 정확도를 평가하는 C-index (concordance index)를 지표로 사용하였다.[11] C-index는 샘플들을 생존시간의 오름차순으로 나열

하고, ‘사건이 관찰된 각 샘플들보다 오래 생존한 샘플들의 개수를 모두 더한 총합’과 ‘샘플들을 예측된 생존시간의 오름차순으로 나열하고, 사건이 관찰된 각 샘플들보다 오래 생존할 것으로 올바르게 예측된 샘플들의 개수를 모두 더한 총합’의 비율 (fraction)로 계산된다. C-index 계산 수식은 다음과 같다.

$$c = \frac{\sum_{i \in U} \left\{ \sum_{T_j > T_i} 1_{f_j > f_i} \right\}}{\sum_{i \in U} \left\{ \sum_{T_j > T_i} 1 \right\}} \quad (3)$$

이 식에서 U는 uncensored 데이터셋이며  $T_i$ 는 샘플  $i$ 의 관측된 생존 시간,  $f_i$ 는 샘플  $i$ 의 예측된 생존 시간이다. 여기서  $1_{a > b}$ 는  $a > b$ 라면 1, 그렇지 않다면 0을 뜻한다. C-index는 0.0 ~ 1.0 사이의 값을 갖는다. 1.0에 가까울수록 정확히 예측한다고 해석하고, 0.5에 가까울수록 무작위로 예측한다고 평가한다.

다중 모드 데이터기반 생존분석 모델의 가능성을 평가하기 위해, Cox 비례위험 모델과 DeepSurv에 대해 임상데이터(Clinical)만을 사용하는 경우와 다중모드 데이터(Clinical+CT+PET)를 함께 사용하는 경우의 C-index값을 비교하였고, 그 결과는 표 1과 같다.

표 2. 다중모드 데이터 사용에 따른 Cox 비례모형과 DeepSurv의 C-index 성능 비교

	Clinical	Clinical+CT+PET
Cox PH	0.7686	0.6564
DeepSurv	0.776575	0.690175

표2에서 DeepSurv모델을 사용할 경우 Cox 비례위험 모델보다 더 높은 C-index값을 얻을 수 있음을 확인할 수 있다. 특히 다중모드 데이터를 사용할 경우, DeepSurv모델이 Cox 비례위험 모델보다 약 5%정도의 성능향상을 얻을 수 있음을 확인하였다.

다만, 현재 다중모드 데이터를 이용할 경우 전반적으로 C-index값이 줄어드는 것을 확인하였다. 이 문제의 해결을 위해서는 영상데이터 특징추출을 위한 효과적인 모델의 연구가 필요하다고 분석 된다.

### 5. 결론

이 논문에서는 임상데이터 뿐만 아니라 영상데이터를 사용하는 생존분석 모델을 제안하고 그 가능성에 대해 평가하였다. 복잡한 특징을 가지는 영상데이터를 활용할 경우, 비선형적 비례위험 모델을 학

습하는 DeepSurv모델이 효과적인 것으로 분석된다. 다만, 다중 모드 데이터를 활용할 경우, 생존분석 성능 향상을 위해서 효과적인 feature 임베딩 기법에 대한 연구가 필요하다.

### Acknowledgment

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 한국연구재단 바이오·의료기술개발사업의 지원을 받아 수행된 연구임(NRF-2019M3E5D1A02067961). 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2017R1A2B4012559).

### 참고문헌

- [1] "사망원인통계결과", 통계청, 2020년
- [2] 박예민, 김명건, et al, "가천의대 길병원에서 진단된 폐암 환자의 임상적 특성", 대한내과학회지, 78권, 2호, 215-221쪽, 2010년
- [3] Chulwoong Choi, Sungwoong Yeom, Hieyong Jeong, Kyungbaek Kim, "CNN based Lung Indexing Method for DICOM CT Image", Proceedings of International Conference on Smart Media and Applications, 2020
- [4] Katzman, J.L., Shaham, U., Cloninger, A. et al. "DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network", BMC Med Res Methodol 18, 24 2018
- [5] Cox D. R, "Regression models and life-tables", Journal of the Royal Statistical Society Series B (Methodological) 34, 187-220p, 1972
- [6] In, Junyong, and Dong Kyu Lee. "Survival analysis: Part I-analysis of time-to-event", Korean journal of anesthesiology, 71권, 3호, 182-191쪽, 2018년
- [7] 강미, Cox의 비례위험모형을 이용한 중소기업의 생존요인분석, 부동산학연구, 15권, 2호, 41-57, 2009
- [8] 이재영 외, Cox's 비례위험 모형을 통한 누적별점의 교통법규 위반 억제효과 분석, 대한교통학회 학술대회지, 257-262, 2015
- [9] 박해상. "DICOM 파일을 이용한 측두골 CT의 3차원 재구성(3D Reconstruction)." 국내석사학위논문 이화여자대학교 대학원, 2010.
- [10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," P roceeding of IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.
- [11] Steck, Harald, et al. "On ranking in survival analysis: Bounds on the concordance index", Advances in neural information processing systems, 1209-1216, 2008