

## Image texture feature를 이용하여 비소세포폐암 전이

### 예측 머신러닝 모델 연구

주혜민<sup>\*,\*,o</sup>, 우상근<sup>\*\*,\*</sup>

\*과학기술연합대학원대학교 방사선종양의과학,

\*\*한국원자력의학원 RI중개연구팀

e-mail: {hmju, skwoo}@kirams.re.kr

## Study of machine learning model for predicting non-small cell lung cancer metastasis using image texture feature

Hye Min Ju<sup>\*,\*,o</sup>, Sang-Keun Woo<sup>\*\*,\*</sup>

\*Radiological and Medico-Oncological Sciences, University of Science and Technology,

\*\*Divison of RI-convergence Research, Korea Institute of Radiology and Medical Sciences

### ● 요약 ●

본 논문에서는  $^{18}\text{F}$ -FDG PET과 CT에서 추출한 영상인자를 이용하여 비소세포폐암의 전이를 예측하는 머신러닝 모델을 생성하였다.  $^{18}\text{F}$ -FDG는 종양의 포도당 대사 시 사용되며 이를 추적하여 환자의 암 세포를 진단하는데 사용되는 의료영상 기법 중 하나이다. PET과 CT 영상에서 추출한 이미지 특징은 종양의 생물학적 특성을 반영하며 해당 ROI로부터 계산되어 정량화된 값이다. 본 연구에서는 환자의 의료영상으로부터 image texture 프질 전이 예측에 있어 유의한 인자인지를 확인하기 위하여 AUC를 계산하고 단변량 분석을 진행하였다. PET과 CT에서 각각 4개(GLRLM\_GLNU, SHAPE\_Compacity only for 3D ROI, SHAPE\_Volume\_vx, SHAPE\_Volume\_mL)와 2개(NGLDM\_Busyness, TLG\_ml)의 image texture feature를 모델의 생성에 사용하였다. 생성된 각 모델의 성능을 평가하기 위해 accuracy와 AUC를 계산하였으며 그 결과 random forest (RF) 모델의 예측 정확도가 가장 높았다. 추출된 PET과 CT image texture feature를 함께 사용하여 모델을 훈련하였을 때가 각각 따로 사용하였을 때 보다 예측 성능이 개선됨을 확인하였다. 추출된 영상인자가 림프절 전이를 나타내는 바이오마커로서의 가능성을 확인할 수 있었으며 이러한 연구 결과를 바탕으로 개인별 의료영상을 기반으로 한 비소세포폐암의 치료 전략을 수립할 수 있을 것이라 기대된다.

**키워드:**  $^{18}\text{F}$ -FDG 양전자 단층 촬영 영상( $^{18}\text{F}$ -FDG PET), 컴퓨터 단층 촬영 영상(computed tomography, CT), 머신 러닝(Machine learning)

## I. Introduction

비소세포폐암은 암중에서 발병률이 높은 질환이며 종양 조직 내 heterogeneity가 매우 높은 특성을 갖고 있다.  $^{18}\text{F}$ -FDG PET은 종양의 포도당 대사 중에 흡수되어 세포에 남아있는 종양 표지자를 추적하여 영상화 한다. 대사 과정에서 종양의 분자생물학적 특성은  $^{18}\text{F}$ -FDG PET 혹은 CT 영상에서 추출한 image texture feature를 통해 확인할 수 있다. 종양 내부의 heterogeneity는 전이 혹은 재발과도 관련이 있는 특성으로 종양의 진단 및 치료반응을 평가하기에 유용한 인자이며 영상에서 추출한 이미지 바이오 마커를 사용하여 정량적인 값을 통한 질병의 예측인자로서 사용할 수 있다. 또한 방사선 영상을 이용하여 비 침습적인 방법으로 개인별 맞춤 치료 전략을 수립할

수 있기 때문에 암의 상태를 이해하고 분석하는 방법으로서 radiomics 기술이 널리 사용되고 있다.

림프절 전이는 비소세포폐암 환자의 생존을 크게 떨어뜨리는 원인 중 하나로 특히 N2 전이가 발생할 경우 N1 전이 보다 예후가 좋지 않다는 연구 결과가 있다. 본 연구에서는 N2 전이를 예측하기 위한 머신러닝 모델을 구축하고 성능을 accuracy와 AUC값을 통해 평가하여 의료 현장에서 image texture feature의 활용 가능성을 제시하고자 한다.

## II. Materials and Methods

### 1. <sup>18</sup>F-FDG PET과 CT영상 준비

머신러닝 네트워크를 학습시키기 위하여 비소세포폐암 환자의 <sup>18</sup>F-FDG PET과 CT영상을 TCIA (<http://doi.org/10.7937/K9/T CIA.2017.7hs46erv>)로부터 다운받아 사용하였다. 환자는 총 66명으로 구성되었으며, 환자의 TNM stage data를 바탕으로 병리학적 단계가 N0인 환자 54명과 N2인 환자는 12명으로 나누어 분석을 진행하였다.

### 2. Image texture feature extraction

비소세포폐암 환자의 <sup>18</sup>F-FDG PET과 CT영상에서 image texture feature를 추출하기 위해 Lifex (version 4.0) 소프트웨어를 사용하였다. 비소세포폐암환자의 종양 부위를 CT 영상을 기준으로 하여 region of interest (ROI)를 설정하였으며, 동일한 ROI를 <sup>18</sup>F-FDG PET영상에도 적용시켜 image texture feature를 계산하였다. ROI 설정 시 semi-automated segmentation 방법이 사용되었으며 이때 설정한 threshold SUV 값은 2.0이었다. <sup>18</sup>F-FDG PET과 CT영상에서 각각 53개의 image texture feature가 계산되었다.

### 3. Image texture feature selection

추출된 image texture feature중에서 전이를 예측함에 있어 유의한 인자를 찾아내기 위하여 R 프로그램에서 AUC를 계산하였다. MedCalc 소프트웨어를 사용하여 receiver operating characteristic (ROC) curve를 그려 sensitivity와 1-specificity를 확인하였다. PET에서는 AUC 값이 0.75이상인 image texture feature가 N2 전이를 예측하는데 인자로서 선택되었으며 CT에서는 AUC가 0.78 이상인 image texture feature가 선택되었다.

### 4. 머신 러닝 성능 평가

전이 결과를 예측하기 위해 머신러닝 모델을 사용하였으며 accuracy와 AUC점수 측면에서 정량적으로 평가되었다. 데이터 불균형을 해결하기 위해 오버샘플링 기법(SMOTE)을 사용하여 오버샘플링 방법을 구현하였다. 모델은 k-fold 검증(K = 10)으로 평가되었다. 전이 관련으로 확인 된 핵심 유전자와 상관관계가 확인 된 image texture feature가 모델을 구축하는데 인자로서 사용되었다. 머신러닝 모델은 전체 환자 중 80%를 대상으로 train을 실시하고 20%의 환자를 대상으로 test를 실시하였다.

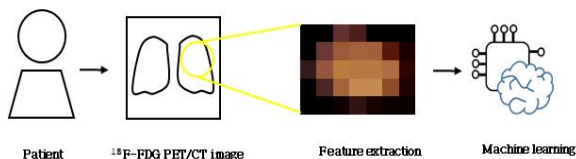


Fig. 1. 머신러닝 모델 생성 모식도

## III. Result

비소세포 폐암환자의 <sup>18</sup>F-FDG PET과 CT영상에서 추출된 전체 image texture feature 중에서 통계적으로 유의미한 값만을 선택하여 머신 러닝 모델을 구축하기 위하여 AUC 값을 계산하였다. PET 영상에서는 AUC 값이 0.75이상인 image texture feature는 통계적으로 유의한 인자로 판단하여 4개의 영상인자(GLRLM\_GLNU, SHAPE\_Compacity only for 3D ROI, SHAPE\_Volume\_vx, SHAPE\_Volume\_mL)가 선택되었다. CT 영상에서는 AUC가 0.78 이상인 2개(NGLDM\_Busyness, GLZLM\_ZLNU)의 image texture feature가 모델의 생성에 사용되었다.

Table 1. AUC value of image texture feature

PET	AUC
GLRLM_GLNU	0.79
SHAPE_Compacity only for 3D ROI	0.77
CT	AUC
NGLDM_Busyness	0.82
GLZLM_ZLNU	0.78

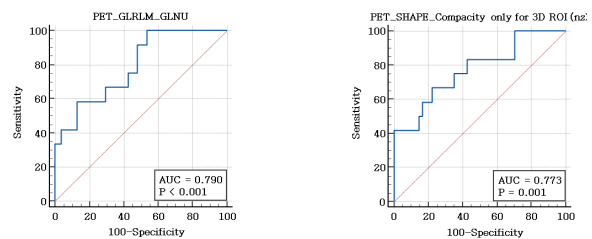


Fig. 2. Image texture feature의 ROC curve

N2 전이를 예측하기 위하여 총 7개의 머신러닝 모델이 사용되었으며, <sup>18</sup>F-FDG PET 과 CT 영상에서 추출된 image texture feature 중에서 AUC를 통해 선택된 영상 인자들이 예측에 사용되었다. <sup>18</sup>F-FDG PET 과 CT 영상에서 추출된 인자를 각각 따로 그리고 함께 사용하여 모델을 구성하였다. <sup>18</sup>F-FDG PET에서 선택한 image texture feature만을 사용한 RF모델의 accuracy와 AUC 값은 0.76과 0.881로 CT에서 선택된 영상인자만을 사용하였을 때 비교하여 감소하였다. <sup>18</sup>F-FDG PET과 CT에서 추출된 image texture feature를 모두 사용하여 RF모델을 구축하였을 때 가장 높은 accuracy와 AUC로 N2를 예측하였다. 총 8개의 머신러닝 모델 중에서 RF모델의 성능이 다른 모델에 비하여 가장 높았다.

Table 2. N2전이 예측을 위한 모델 성능 평가

	CT (n= 2)		<sup>18</sup> F-FDG PET/CT (n = 6)	
	accuracy	AUC	accuracy	AUC
LR	0.674	0.827	0.683	0.812
LDA	0.702	0.832	0.683	0.798
KNN	0.706	0.833	0.769	0.816
CART	0.843	0.848	0.742	0.722
NB	0.655	0.71	0.684	0.81
SVM	0.637	0.796	0.674	0.804
RF	0.853	0.898	0.816	0.898

#### IV. Conclusions

본 연구에서는 비소세포폐암의 N2 전이를 예측하기 위하여 <sup>18</sup>F-FDG PET과 CT에서 추출한 image texture feature를 사용한 머신러닝 모델을 구축하고, 모델 성능을 평가하였다. 추출된 image texture feature는 바이오 마커로서 높은 가능성을 보여주었다. 전이를 예측하기 위해 생성된 8개의 머신러닝 모델 중 RF모델이 가장 높은 성능을 나타내었으며 이 연구를 바탕으로 암환자의 정확한 림프절 전이를 예측하고 조기에 치료 전략을 수립하는데 있어서 큰 도움이 될 수 있을 거라 기대된다.

#### REFERENCES

[1] Chen, Z.; Fillmore, C.M.; Hammerman, P.S.; Kim, C.F.; Wong, K.K. Non-small-cell lung cancers: A heterogeneous set of diseases. *Nat. Rev. Cancer* **2014**, *14*, 535-546.

[2] Cai, M.; Wang, Z.; Zhang, J.; Zhou, H.; Jin, L.; Bai, R.; Weng, Y.J.C.P. Adam17, a target of Mir-326, promotes emt-induced cells invasion in lung adenocarcinoma. *Cell. Physiol. Biochem.* **2015**, *36*, 1175-1185.

[3] O'Brien, B.; van der Putten, W. Quantification of risk-benefit in interventional radiology. *Radiat. Prot. Dosim.* **2008**, *129*, 59-62.

[4] Yoon, H.J.; Sohn, I.; Cho, J.H.; Lee, H.Y.; Kim, J.H.; Choi, Y.L.; Kim, H.; Lee, G.; Lee, K.S.; Kim, J. Decoding tumor phenotypes for ALK, ROS1, and RET fusions in lung adenocarcinoma using a radiomics approach. *Medicine* **2015**, *94*, e1753.

[5] Bianconi, F.; Palumbo, I.; Fravolini, M.L.; Chiari, R.; Minestrini, M.; Brunese, L.; Palumbo, B. Texture analysis on [<sup>18</sup>F]FDG PET/CT in non-small-cell lung cancer: Correlations between PET features, CT features, and histological types. *Mol. Imaging Biol.* **2019**, *21*, 1200-

1209.

[6] Andersen, M.B.; Harders, S.W.; Ganeshan, B.; Thygesen, J.; Torp Madsen, H.H.; Rasmussen, F. CT texture analysis can help differentiate between malignant and benign lymph nodes in the mediastinum in patients suspected for lung cancer. *Acta Radiol.* **2016**, *57*, 669-676.