Vision transformer using FDG PET for survival prediction in lung cancer patients

Sae-Ryung Kang1, Seungwon Oh2, Jung-Joon Min1, Hee-Seung Bom1, In-Jae Oh3, Soo-Hyung Kin4, Guee-Sang Lee4, Hyung-Jeong Yang4, Min-soo Kim5, Seong-Whan Lee2,6

1Department of Nuclear Medicine, Chonnam National University Medical School and Hwasun Hospital, Jeonnam, Republic of Korea2Department of Artificial Intelligence, Korea University, Seoul, Republic of Korea3Deparment of Internal Medicine, Chonnam National University Medical School and Hwasun Hospital, Jeonnam, Republic of Korea4Department of Artificial Intelligence Convergence, Chonnam National University, Gwangju, Republic of Korea5Department of Mathematics & Statistics, Chonnam National University, Gwangju, Republic of Korea6Department of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea.

Purpose: Accurate survival prediction is essential to guide treatment in lung cancer patients. F-18 fluorodeoxyglucose positron emission tomography (FDG PET) imaging has been used for survival prediction via radiomics features. We aimed to improve the performance of survival prediction using vision transformer (ViT) based deep learning model with FDG PET in lung cancer patients.

Methods: We proposed a ViT based survival prediction model using multimodal approach; FDG PET images, tumor volume mask PET images, and clinical data (age, sex, TNM stage, histology, smoking history). 3D PET images and tumor volume mask images were divided into multiple patches. Each patch image was projected onto axial, sagittal, and coronal planes with two different ways; maximum intensity projection and sum intensity projection. The latent variables extracted through convolutional neural network (CNN) from projected PET and tumor volume mask images were combined and vectorized by transformer encoder by adding positional encoding. Survival prediction using DeepSurv was performed by combining clinical features and vectorized image features. We evaluated the performance of the proposed model via concordance index and Brier score under a five-fold cross-validation in 400 consecutive lung cancer patients (343 non-small cell lung cancer, 57 small cell lung cancer) who were diagnosed between 2017 and 2020 at Chonnam National University Hwasun Hospital.

Results: The proposed ViT based survival prediction model showed better performance than conventional Cox Proportional Hazards (CPH) model in terms of evaluation metrics of concordance index (0.779±0.022 vs. 0.727±0.027, $p < 0.05$) and Brier score (0.130±0.002 vs. 0.154±0.009, $p < 0.01$).

Conclusions: Our proposed deep learning based survival prediction model showed improved

performance using FDG PET images with tumor volume masks and clinical data.