PET-Based Deep-Learning Model for Predicting Prognosis of Patients With Non-Small Cell Lung Cancer

SEUNGWON OH\textsuperscript{1}, JAENA IM\textsuperscript{1}, SAE-RYUNG KANG\textsuperscript{2}, IN-JAE OH\textsuperscript{3}, AND MIN-SOO KIM\textsuperscript{1}

\textsuperscript{1}Department of Mathematics and Statistics, Chonnam National University, Buk-gu, Gwangju 61186, Republic of Korea
\textsuperscript{2}Department of Nuclear Medicine, Chonnam National University Medical School and Hwasun Hospital, Jeonnam 58128, South Korea
\textsuperscript{3}Department of Internal Medicine, Chonnam National University Medical School and Hwasun Hospital, Jeonnam 58128, South Korea

Corresponding authors: In-Jae Oh (droij@chonnam.ac.kr) and Min-Soo Kim (kimms@chonnam.ac.kr)

This work was supported in part by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean Government (MSIT) under Grant NRF-2019M3E5D1A02067961, and in part by Korea Health Technology Research and Development Project through Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health and Welfare, Republic of Korea, under Grant HR20C00021.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board of Chonnam National University Hwasun Hospital under Application No. CNUHH-2019-194.

ABSTRACT

Despite recent advances in precision medicine, lung cancer remains the leading cause of cancer-related mortality worldwide. To determine the prognosis of non-small cell lung cancer (NSCLC), which accounts for 85% of lung cancer, comprehensive analysis of various clinical factors are necessary. Artificial intelligence can help physician quickly identify key information from the vast amount of medical information including positron emission tomography (PET) scan. In this study, we compared image feature-extraction models and survival estimation models to determine an optimal model that effectively extracts features related to survival time. We collected PET image data of 2,685 patients who were diagnosed with NSCLC and received treatment at the Chonnam National University Hwasun Hospital in South Korea over a period of seven years. We compared four convolution neural network models, DenseNet, NFNet, EfficientNet, and ResNet, and two survival estimation models, CoxPH and CoxCC. The best model was determined based on criteria such as C-index, mean absolute error (MAE), classification accuracy for survival status, and learning time. The results show that DenseNet combined with CoxPH delivers superior performance for most of the criteria. In particular, the MAE for this combination was very low (391.50 days), and the model predicted survival days well; the five-year classification accuracy, which can indicate a cure for cancer, was high (95%). Extracted features were visualized using Score-CAM; thus, the learning process of the model could be understood without requiring expert knowledge of PET. In addition, the learning time for this model was short.

INDEX TERMS

Convolutional neural network, CoxPH, CoxCC, non-small cell lung cancer prognosis, Score-CAM, positron emission tomography, F-18 fluorodeoxyglucose.

I. INTRODUCTION

Lung cancer is the most commonly diagnosed cancer in both men and women worldwide. Accounting for 11.6% of all cancer diagnoses, lung cancer has the highest diagnosis rate of all cancers; the mortality rate is also the highest at 18.4% of all cancer deaths [1]. In addition, non-small-cell lung cancer (NSCLC) accounts for 85% of all lung tumors detected [2]. Only a small number of patients with lung cancer are diagnosed through screening, and most are diagnosed through the presentation of several symptoms [3], [4]. However, most symptoms of lung cancer only occur once the disease has progressed to some extent, making treatment difficult. Currently, the mortality risk of patients is assessed through the examination of many factors, such as age, smoking history, previous cancer history, family history, and radiologic factors, such as tumor size and shape. There are various treatment strategies according to tumor stage, location, histology and genetic
alteration, and many hospitals constitute multidisciplinary teams to decide best treatment option. In addition, novel drugs, medical data, papers and treatment guidelines for lung cancer are rapidly developed. Therefore, a time-consuming and labor-intensive process must be followed to treat and analyze the prognosis of lung cancer patients because various factors must be considered.

Various radiologic factors are used to determine the staging of patients with NSCLC. The most representative method is the TNM staging system of the American Joint Committee on Cancer (AJCC). The TNM staging system uses three values to describe the progression of the cancer: T indicates the size of the tumor, N indicates metastasis to the lymph nodes, and M indicates metastasis to distant organs. Higher numbers represent more advanced stages of cancer. A criterion for estimating residual life using the TNM staging system was proposed [5]. However, it is difficult to reflect the individual characteristics of patients using only staging. Therefore, studies have been conducted to investigate the use of statistical models [6]–[8].

In traditional survival analysis studies, the Cox proportional hazards model (CPH) is the most-used semi-parametric model [9], [10]. CPH is a regression model, and it estimates the hazard ratio by linearly combining several covariates. An advantage of the regression model is that each variable is easy to interpret, but there is a limit to expressing the nonlinear relationship between covariates. In addition, it disadvantage is its inability to learn unstructured data such as image data.

Several studies used deep learning to overcome the disadvantage of linear models [11]–[13]. Cheerla et al. constructed a multimodal neural-network-based model to predict the survival of patients with 20 different cancer types using clinical data, mRNA expression data, microRNA expression data, and histopathology whole slide images (WSIs) [11]. Zhu et al. [13] proposed a deep convolutional neural network that uses pathological images for survival analysis (DeepConvSurv). Coudray et al. [12] trained a deep convolutional neural network (inception v3) on WSIs obtained from The Cancer Genome Atlas to classify them accurately and automatically into LUAD, LUSC, or normal lung tissue.

Positron emission tomography (PET) imaging is a functional imaging technique that can visualize and measure the disease status semiquantitatively using tracers labeled with radioactive isotopes [14], [15]. F-18 fluorodeoxyglucose (FDG) PET is routinely obtained to detect tumors in whole body of lung cancer patients. FDG PET has been actively investigated to predict the prognosis of cancer patients because they provide information of tumor pathophysiology reflecting prognosis [16], [17]. However, extraction of handcrafted features from the PET images for predicting prognosis requires professional knowledge of expert and is time-consuming. Recently, deep convolutional neural networks (CNNs) have been investigated to overcome these limitations, but considerable amount of computing resources and time are needed for predicting prognosis using hundreds of images from whole body PET of each patient.

Therefore, we aimed to predict the individual survival distribution and time in NSCLC patients using only maximal intensity projection (MIP) image which is a 2 dimensional image used routinely by physicians in hospitals without other factors such as clinical stage. For feature extraction from PET images, the CNN models that performed well in the ImageNet Large Scale Visual Recognition Challenge were used. In addition, combinations of the CNN models and survival estimation models were used to estimate the hazard ratio of patients. Each model was evaluated to determine the best model, which was finally selected based on various metrics such as classification accuracy for survival status, C-index, mean absolute error (MAE), and learning time.

II. MATERIALS AND METHODS
A. CLINICAL AND PET DATA
The PET image data and overall survival time were collected in NSCLC patients diagnosed and treated at Chonnam National University Hwasun Hospital (CNUHH) in South Korea from January 2011 to December 2017 Overall survival time was calculated as the period from the date of
TABLE 1. Patients' clinical characteristics in training and test data sets.

| Characteristics                     | Number of patients | Training set | Test set | P-value |
|-------------------------------------|--------------------|--------------|----------|---------|
| Survival time (days)                | 2,685              | 819.96±756.773 | 813.37±758.458 | 0.857   |
| Age                                 | 2,685              | 67.97±9.705   | 67.86±9.327 | 0.800   |
| Smoking amount (pack-years)         | 2,685              | 29.18±27.157  | 29.96±27.553 | 0.550   |
| Dead status                         |                    |              |          |         |
| Censored                            | 828                | 662 (80.0%)  | 166 (20.0%) | 0.966   |
| Death                               | 1,857              | 1,485 (80.0%) | 372 (20.0%) |         |
| Smoking                             |                    |              |          |         |
| Non-smoking                         | 826                | 664 (80.4%)  | 162 (19.6%) | 0.753   |
| Smoking                             | 1,859              | 1,483 (79.8%) | 375 (20.2%) |         |
| Gender                              |                    |              |          |         |
| Female                              | 681                | 546 (80.2%)  | 135 (19.8%) | 0.916   |
| Male                                | 2,004              | 1,601 (79.9%)| 403 (20.1%)|         |
| Time interval from PET to diagnosis (days) | 2,685   | 4.38±10.252 | 4.64±13.028 | 0.669   |
| Total                               | 2,685              | 2,147        | 538      |         |

1 the time interval from the date of the PET scan to the date of clinical diagnosis.

diagnosis to the date of death. All patients underwent FDG PET/CT scans prior to initiation of treatment. To test generalization, the PET image data were derived from two types of PET/CT scanners: Discovery ST (GE Medical Systems, Milwaukee, WI, USA) and Discovery 600 (GE Medical Systems, Milwaukee, WI, USA) [18], [19]. The PET scans were performed according to standardized imaging protocols at CNUHH. Patients were instructed to fast for at least 6 h before injection of F-18 FDG, and the serum glucose level were found to be 8.3 mmol/l or less. PET image acquisition for torso scanning was begun at about 1 h after the intravenous injection of 7.4 MBq/kg body weight of F-18 FDG. A low-dose CT scan was performed for attenuation correction from the head to the thigh. CT was performed using the following parameters: 120 kV, 10–130 mA, rotation time 0.7 s, field of view (FOV) 50 cm, and slice thickness 3.75 mm. Immediately after the CT scan, PET was performed with 15.7 cm axial FOV acquired in 2D mode with 150 s/bed (Discovery ST) and 120 sec/bed (Discovery 600) position covering the same field of CT. PET images were reconstructed iteratively with an ordered subset expectation maximization (OSEM) algorithm. The OSEM reconstruction algorithm was evaluated for different reconstruction settings in phantom studies and clinical settings. The algorithm was optimized (2 iterations, 16 subsets, and 6.4 filter cutoff) to obtain the best quality of reconstructed images. This study was approved by the Institutional Review Board of our institution (CNUHH-2019-194).

The log-rank test of each variable for patient property was statistically significant under the significance level of 0.05. This means that there is a difference in the survival function according to the patient property. In particular, the survival probability decreases when the patient properties are- old, male, and smoking (Fig 2).

Training image classification models using 3D images requires considerable computing resources and time. In addition, it is difficult to visualize the 3D images used for model interpretation after training. Maximum intensity projection (MIP) is a simple 3D visualization tool that can be used to display 3D PET datasets [20]. MIP projects the 3D image data onto a plane using the maximum intensity value of the voxels. In this study, MIP projected the 3D PET data onto a coronal plane. MIP images are not threshold-dependent and preserve attenuation information. All the MIP images of the NSCLC patients had a width of 128 and a height of 427 (Fig 1).
Resizing, cropping, and padding are representative methods of processing images having different sizes. In the resizing method, the aspect ratio of the image is adjusted. The cropping method crops an image to the smallest image size. In the padding method, the empty space is filled with the same pixel value as the background based on the largest image through padding. Because this study needs to extract features such as the size and shape of lung cancer tumors using MIP images, there is a risk of distortion when using the resizing method, which changes the ratio. In addition, there is a problem in that it is difficult to reflect the information of whole-body images of male patients when cropping is performed based on the minimum length. Therefore, for image standardization, padding was performed based on the largest height. In addition, CNNs can extract local features, and the padded part has less effect on the model [21]. Therefore, we converted all the images through padding to a 427-height format, which has the most slices (Fig 3).

Five-fold cross validation was performed for model training and generalization. The split of the data was maintained at the same censoring rate using stratified sampling based on the censored data. For the similarity of the divided each data, the continuous variable was tested through the t test, and the categorical features were tested through the chi-square test. The results revealed that the training and the test sets were not statistically different under the significance level of 0.05, showing that the two datasets were balanced (Table 1).

### B. CNN MODEL

To extract the features related to the hazard ratio from the PET images, the best CNN model out of four was selected. CNN is a deep learning model for data that has a grid pattern, such as images. It is a mathematical construct that is typically composed of three layers, including convolution, pooling, and fully connected layers. Stacking layers are an important factor in improving performance, but degradation problems may occur due to vanishing/exploding gradients. Although these limitations can be modified through weight adjustments such as weight initialization and batch normalization, the following models can overcome in their own way (Fig 4). Thus, we compared the following models that exhibited a satisfactory performance on the ImageNet dataset [22].

ResNet is an algorithm developed by Microsoft, and its model won the ILSVRC in 2015. It contains shortcut connections which turn the network into its counterpart residual version was inserted. The deep residual learning framework allows few stacked layers to fit residual mapping instead of expecting these layers to directly fit a desired underlying mapping. Residual learning is adopted for every few stacked layers. The shortcut connections introduce neither extra parameters nor computation complexities [23]. The residual block used by ResNet is later used variously for other models [24].

DenseNet simplifies the connectivity pattern between layers to keep increasing the depth of deep convolutional networks. The model solves the problem of ensuring maximum information flow by simply connecting every layer directly with each other. Instead of drawing representational power from extremely deep or wide architectures, DenseNet exploits the potential of the network through feature reuse [25].

EfficientNet is a new compound scaling method, which uses a compound coefficient to uniformly scale network width, depth, and resolution in a principled way. To better demonstrate the effectiveness of the scaling method, this new mobile size baseline called EfficientNet was developed [26]. Most of the papers that achieved SOTA performance on ImageNet data used EfficientNet [27]–[30].

Normalizer-Free ResNet called NFNets is a family of modified ResNets that achieves competitive accuracies without batch normalization. It proposes adaptive gradient clipping (AGC), which clips gradients based on the unit-wise ratio of gradient norms to parameter norms. To train deep ResNets without normalization, it is crucial to suppress the scale of the activations on the residual branch. To achieve this, NFNets uses two scalers to scale the activations at the start and end of the residual branch [31].

### C. SURVIVAL MODEL

We compared the time-to-event predictions of the following survival estimation models:

- **DeepSurv**: a Cox proportional hazards model.
- **CoxCC**: an extension of the CPH model.

The DeepSurv model can be applied to a neural network using negative log partial likelihood with L2 regularization loss [32]. Survival analysis studies for cancer using DeepSurv were conducted previously [33], [34]. In particular, Matsuo et al. [34] compared various Cox and deep-learning models with clinical factor data for female patients, and another study showed slightly improved results on the MAE evaluation index compared to those of existing machine learning models. CoxCC is a semi-parametric form of the Cox model, in which the relative risk function is changed over time [35].
TABLE 2. Comparison of CNN and survival estimation models in five-fold.

| CNN model (#Params) | Survival model | C-index | MAE (days) | Accuracy (2-year survival status) | Accuracy (5-year survival status) | Learning time (seconds) |
|---------------------|----------------|---------|------------|-----------------------------------|-----------------------------------|------------------------|
| CPH                 | CoxCC          | 0.61±0.013 | 780.60±32.701 | 0.33±0.007                        | 0.89±0.012                        | -                      |
| DenseNet (6.9M)     | CoxCC          | 0.44±0.045 | 1065.68±137.418 | 0.24±0.014                        | 0.86±0.085                        | 439.20±5.982           |
|                     | CoxPH          | 0.71±0.009 | 391.50±29.222  | 0.71±0.027                        | 0.95±0.009                        | 434.17±6.655           |
| NIPNet (68.4M)      | CoxCC          | 0.68±0.016 | 667.04±71.927  | 0.64±0.020                        | 0.83±0.023                        | 1197.86±2.495          |
|                     | CoxPH          | 0.68±0.013 | 561.76±71.682  | 0.68±0.015                        | 0.89±0.029                        | 1207.59±8.877          |
| EfficientNet (4.0M) | CoxCC          | 0.50±0.006 | 2230.62±829.136 | 0.28±0.054                        | 0.26±0.266                        | 249.34±2.054           |
|                     | CoxPH          | 0.54±0.021 | 892.55±389.544  | 0.47±0.131                        | 0.75±0.210                        | 248.26±4.314           |
| ResNet (42.5M)      | CoxCC          | 0.51±0.020 | 2192.21±672.858  | 0.27±0.050                        | 0.25±0.237                        | 261.72±6.640           |
|                     | CoxPH          | 0.53±0.022 | 1253.08±68.016  | 0.47±0.032                        | 0.65±0.023                        | 279.94±7.360           |

Figure 5. Comparison of model performance using MAE (a), C-index (b), and classification accuracy for 5-year survival status (c).

III. EVALUATION CRITERIA

To evaluate the performance of the models, the following metrics were used.

A. C-INDEX

To evaluate the performance of the survival estimation model, we need to consider the relative risk of the event. The most common evaluation method is the concordance index (C-index), which indicates the accuracy of the ranking of the predicted time [36]–[38]. Perfect concordance is represented by 1.0, and a poor prediction is represented by 0.5.

\[ C-index = \frac{\# \text{correctly ordered pairs}}{\# \text{all possible ranking pairs}} \]

B. MAE

Since the C-index does not reflect information on the prediction time but only indicates the accuracy of the ranking of predicted time, a high score does not necessarily indicate the accuracy of the time prediction of the model. The MAE is the average of the difference between the predicted residual life and the survival time among patients who actually died (\( N_{E=1} \)). The individual survival function can be estimated using the model’s output hazard ratio, and the estimated survival function is a non-parametric survival function for individual patients. We used the median residual life (MRL) method, which estimates the non-parametric residual life for the survival time at which the survival probability falls below 0.5 [39].

\[ \hat{y}_i = \delta(0.5), \]

\[ MAE = \frac{1}{N_{E=1}} \sum_{i=1}^{n} N_{E=1} \times |y_i - \hat{y}_i| \]

C. CLASSIFICATION ACCURACY

MAE faces the problem that it converges on the average survival time. Therefore, to compare the accuracy of the range for survival time, we compared MAE using year classification. The residual life estimation was used to assess the accuracy of the patient’s true survival time. Accuracy was evaluated based on a binary classification of two and five years.

D. EFFICIENCY

In general, deep-learning performance improves when more layers or nodes are used. However, excessive parameters can result in overestimation and a lack of computing resources. Therefore, a short learning time and a small number of parameters can be considered as indicators of a good model because computer resources can be saved.
S. Oh et al.: PET-Based Deep-Learning Model for Predicting Prognosis of Patients With NSCLC

IV. RESULT

All models were simulated under the same conditions; learning rate and epoch are 1e-5, 50 respectively. The CPH model, which uses only patients’ properties, shows poor performance in all evaluations because other clinical information such as clinical stage is not reflected. After evaluating the performance of the two survival estimation models and the four CNN models, DenseNet using CoxPH exhibited a satisfactory performance for most of the criteria. DenseNet using CoxPH showed the best performance in terms of MAE with an error of 391.50 days. In terms of C-index, the performance of DenseNet with CoxPH was similar to that of NFNet using both survival estimation models. DenseNet using CoxPH showed the best performance in both classification of 2-year and 5-year survival status. In particular, the five-year classification accuracy, which can indicate a cure for cancer, was very high at 95%. When comparing learning time, EfficientNet and ResNet showed satisfactory performance, but they did not perform well in terms of the other indicators. Finally, the model using DenseNet and CoxPH exhibited the best overall performance (Table 2, Fig 5).

Deep learning exhibits excellent performance, but it is difficult to interpret. To solve this problem, visualization using Class Activation Map (CAM) for image data was proposed [40]. In this study, the features were visualized using Score-CAM [41]. The features of the DenseNet model using CoxPH were visualized to show that parts of the PET image that affected the risk ratio. Fig 6 shows the process of DenseNet extracting features from a low to high. Fig 6 (a) shows the original MIP image. Fig 5 (b) shows the low-level features in the first block of DenseNet. Low-level features show the high signal intensity sites such as the brain, tumors, urinary bladder, and tracer injection site. Fig 6 (c) shows the selected features related to survival time, excluding non-critical features such as the brain and injection site. Fig 5 (d) indicates the detection of more important features. Fig 6 (e) shows that the model focuses on the most important features on lung tumors.

Fig 7 shows the Kaplan-Meier plot that was used to compare the distribution of true survival time and predicted survival time combined with CoxPH [42]. The similarity evaluation of survival distribution was not statistically significant under the significance level of 0.05 when using the log-rank test except EfficientNet using CoxPH [43]. Therefore, distributions of the true and predicted survival time three CNN model combined with CoxPH can be considered as similar.

In addition, the survival function of each NSCLC patient can be estimated through the predicted hazard ratio. Fig 8 illustrates the results of estimating survival curves for patient A and B. For each CNN model, only the results were used when combined with CoxPH. Fig 8 (a) is the predicted survival curve of a 74-year-old male who smokes.

**FIGURE 6.** Score-CAM feature map visualization of each block produced by the DenseNet model.

**FIGURE 7.** Kaplan-Meier plot of the test set.

**FIGURE 8.** Estimated individual survival plot and survival time.
Survival time (days) was as short as 124 days, and all four models showed little error in survival days. However, Fig 8 (b), shows that, in the case of a 72-year-old female non-smoker, the survival time was 1,131 days, which is longer than the median survival time (608 days). In this case, the other three models all overestimate the survival time, while DenseNet has a value of 1,246.05 (days), which is similar to the ground truth. In estimating individual survival curves and predicting residual life, DenseNet and CoxPH show better results than other models.

V. DISCUSSION

For the better treatment outcome, several clinical decision-support systems such as Watson for Oncology are being applied in several cancer fields [44]. And deep learning has the potential to be extremely useful in medicine, particularly in the interpretation of medical images such as computed tomography, PET and histopathological slides [45], [46]. Recently, FDG PET has been promising imaging tool to predict patient outcomes. However, PET data, as a large three-dimensional image data, is difficult to handle it. Our study attempted to perform survival prediction using single MIP PET image in NSCLC patient individually to help doctors and patients make the best decisions.

In a previous study, the C-index of the CPH model using various patient information such as gender, age, clinical stage, histological diagnosis, mutation, tobacco use, weight loss, ECOG performance status, and respiratory comorbidities was 0.72 (validation cohort 1) and 0.71 (validation Cohort 2) [47]. To predict the patient outcome in lung cancer, it is a current usual practice to obtain various medical test results, such as blood test, imaging tests, and histopathological tests, and patient characteristics information.

MIP converts a 3-dimensional image with a large volume into a 2-dimensional image. Therefore, coronal MIP image has low sensitivity in detecting tumors, and it is difficult to localize tumors in 3 dimensions. However, in a previous study, coronal MIP PET input improved the performance in localizing and classifying uptake patterns of whole-body PET images in patients with lung cancer, compared to routine axial PET/CT input [48]. It was considered that MIP images enable readers to capture uptake patterns in the whole body at a first glance and recognize tumors more easily. In addition, the CNN model can have great advantages in visualization using CAM for 2-dimensional images. In future research, it is necessary to use rotated images to reduce depth information and loss of information.

In this study, features were extracted using the basic form of each CNN models widely used in the image processing field, and the survival distribution was estimated using two deep learning-based survival models. The CNN structures that effectively extract information on lung cancer tumors are concluded to be DenseNet and NFNet, and the models that showed the best performance were the models using DenseNet and CoxPH. However, DenseNet showed poor performance when CoxCC was used because it was sensitive to differences in survival models, whereas NFNet showed consistent performance regardless of survival models for both CoxPH and CoxCC. Clinical data using multi-layer perceptron (MLP) in the previous study showed similar performance to CoxPH and CoxCC [49]. Therefore, it is concluded that NFNet has superior stability of extracted features. Although ResNet showed good performance in previous studies, it was difficult to learn how to increase the depth of the layer with the current data of 2,685 people because of the large number of parameters (42.5M). In addition, EfficientNet was expected to learn well with the fewest parameters (4.0M). However, the performance deteriorated because it could not find an appropriate compound scaling coefficient owing to a small amount of training data unlike ImageNet. In addition, this study has limitations in feature extraction because only the most basic form of each model is used. Another limitation is that the training and validation data of the model in this study were all from a single center, and the predictive performance of the models was not externally verified.

In this study, the prediction of individual survival distribution and residual life of NSCLS patient was performed using a deep-learning approach with only a MIP image of PET. It is a cost-effective and easy method to use as a survival prediction tool in real-world practice. Moreover, feature extraction was performed using a CNN model that showed excellent performance on ImageNet, and the learning process of the model was visualized using Score-CAM. Thus, the learning process of the model could be understood without professional knowledge regarding PET, and the learning time was shortened. However, the proposed model used relatively small amount of information from PET images, compared to the actual amount of medical information obtained in real-world practice and the performance was not satisfactory to use in real world. It might be necessary to use raw data of whole body PET images and multimodal medical data that can consider various factors for more accurate prediction of survival of patients in future studies.

REFERENCES

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, “Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” CA Cancer J. Clin., vol. 68, no. 6, pp. 394–424, 2018.

[2] A. Jemal, R. Siegel, E. Ward, T. Murray, J. Xu, C. Smigal, and M. J. Thun, “Cancer statistics, 2006,” CA Cancer J. Clin., vol. 56, no. 2, p. 106, 2006.

[3] K.-H. In, “Lung cancer patients who are asymptomatic at diagnosis show favorable prognosis: A Korean lung cancer registry study,” Lung Cancer, vol. 64, no. 2, p. 232–237, May 2009.

[4] J. Y. Park and S. H. Jang, “Epidemiology of lung cancer in Korea: Recent trends,” Tuberculosis Respiratory Diseases, vol. 79, no. 2, p. 58, 2016.

[5] P. Goldstraw, K. Chansky, J. Crowley, R. Rami-Porta, H. Asamura, W. E. Eberhardt, A. G. Nicholson, P. Groome, A. Mitchell, and V. Bolejack, “The IASLC lung cancer staging project: Proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM classification for lung cancer,” J. Thoracic Oncol., vol. 11, no. 1, pp. 39–51, 2016.

[6] L. Han, R. Kong, D.-D. Yin, E.-B. Zhang, T.-P. Xu, W. De, and Y.-Q. Shu, “Low expression of long noncoding RNA GAS6-AS1 predicts a poor prognosis in patients with NSCLC,” Med. Oncol., vol. 30, no. 4, p. 694, Dec. 2013.
[49] Y. Wu, J. Ma, X. Huang, S. Ho Ling, and S. W. Su, “DeepMMSA: A novel multimodal deep learning method for non-small cell lung cancer survival analysis,” 2021, arXiv:2106.06744. [Online]. Available: http://arxiv.org/abs/2106.06744

SEUNGWON OH received the B.S. and M.S. degrees in statistics from Chonnam National University, South Korea, in 2016 and 2018, respectively, where he is currently pursuing the Ph.D. degree. His research interests include machine learning algorithm and artificial intelligence methodologies.

JAENA IM received the B.S. degree in statistics from Chonnam National University, South Korea, in 2020, where she is currently pursuing the M.S. degree. Her research interests include survival analysis, machine learning algorithm, and deep learning.

SAE-RYUNG KANG received the M.D., M.S., and Ph.D. degrees from Chonnam National University Medical School, South Korea. She is currently serving as a Clinical Assistant Professor for the Department of Nuclear Medicine, Chonnam National University Hwasun Hospital. Her research interests include molecular imaging, radiomics, and artificial intelligence.

IN-JAE OH received the M.D., M.S., and Ph.D. degrees from Chonnam National University, South Korea. He is currently a Professor with the Department of Internal Medicine, Chonnam National University Medical School and Hwasun Hospital. His research interests include molecular diagnosis and translational research of lung cancer.

MIN-SOO KIM received the B.S., M.S., and Ph.D. degrees in computer science and statistics from Chonnam National University, Gwangju, South Korea, in 1994, 1996, and 2000, respectively. From 2003 to 2005, he was a Postdoctoral Fellow at Korea Advanced Institute of Science and Technology(KAIST). In 2007, he joined the faculty of the Department of Statistics, Chonnam National University, where he has been serving as the Chairperson of the Big Data Center, since 2015. His research interests include pattern recognition, wavelet, artificial intelligence and multivariate statistics, and analysis of big data according to climate changes.