Predicting Lung Cancer Patient Survival with Discrete-Time Survival Wide-and-Deep Learning

Steven Z. Wang

Solon High School, Solon, OH 44139, United States

* Corresponding author email: stevenzw8@gmail.com

Abstract

Lung cancer remains a leading cause of cancer-related deaths worldwide, necessitating improved prognostic tools for personalized treatment strategies. This study introduces a novel discrete-time survival wide-and-deep learning framework for predicting lung cancer patient survival using multi-source data. Our approach synergizes discrete-time survival models with wide-and-deep learning architectures to effectively integrate clinical variables, radiomic features, and raw CT images. The proposed model comprises a wide component that captures cross-product relationships with some linear transformations among clinical and radiomic features and a deep component utilizing 3D convolutional neural networks (CNNs) to extract high-level semantic information from raw CT images. This architecture allows for both recognition of key feature interactions and generalization to unseen patterns, providing promising prediction capability in medical fields. In practice, we applied our model to a dataset of 388 non-small cell lung cancer patients, comparing its performance against a random survival forest baseline model. The wide-and-deep model achieved superior predictive accuracy. Compared to the random forest model, the wide-and-deep model demonstrated a much lower integrated Brier score on the test set. Our findings highlight the potential of this innovative approach in enhancing prognostic accuracy in lung cancer prediction, contributing to the growing intersection of artificial intelligence, radiomics, and personalized medicine.

Keywords

Deep learning; Survival analysis; Prediction; Lung Cancer; Medical Image

Introduction

Lung cancer remains one of the most devastating malignancies worldwide, accounting for a significant proportion of cancer-related deaths. In the United States alone, it is responsible for more cancer fatalities than any other type and causes more deaths than colorectal, breast, and prostate cancers combined. The aggressive nature of lung cancer, coupled with its often asymptomatic early stages, presents a formidable challenge in oncology. Most patients are diagnosed at advanced stages when the cancer has already metastasized, severely limiting treatment efficacy and reducing survival rates. Conversely, early detection significantly improves the chances of curative treatment, underscoring the critical importance of timely diagnosis and accurate prognosis.[1]

Recent advancements in medical imaging, particularly low-dose computed tomography (LDCT), have shown promise in early lung cancer detection. The National Lung Screening Trial (NLST) demonstrated that three annual rounds of LDCT screening in high-risk populations reduced lung cancer mortality by 20% compared to chest radiography screening over a seven-year period.[2] This landmark study has catalyzed the implementation of LDCT-based lung cancer screening programs across the United States, marking a significant step forward in early detection efforts.

Parallel to these developments in medical imaging research, the field of artificial intelligence (AI) and machine learning (ML) has been rapidly evolving, offering new avenues for enhancing medical diagnostics and prognostics. In particular, deep learning (DL) algorithms have shown remarkable

capabilities in analyzing complex, high-dimensional data such as medical images.[3,4] The application of these advanced computational methods to oncology has given rise to the field of radiomics - the high-throughput extraction of quantitative features from radiographic images. Radiomics allows for the comprehensive quantification of tumor phenotypes by applying a large number of quantitative image features.[5,6] These features can capture subtle patterns of tumor heterogeneity, shape, and texture that may be imperceptible to the human eye but could hold significant prognostic value. Previous studies have demonstrated that radiomic features extracted from CT images can provide valuable insights into tumor biology, treatment response, and patient outcomes in lung cancer.

Although there are many different types of valuable data recording in lung cancer research, it is still a challenging and meaningful topic to accurately predict the survival time of lung cancer patients. In recent years, several deep learning models have been proposed for survival analysis, including both continuous-time and discrete-time approaches. These models build upon the foundation of traditional survival analysis techniques while leveraging the power of neural networks. DeepSurv[7] adapted the conventional Cox proportional hazards regression model to a deep neural network architecture, enabling more complex relationships between covariates and survival outcomes to be captured. It trains the network by setting the objective function to be the average negative log (Cox) partial likelihood with regularization. Kvamme et al. [8] proposed a new method for time-to-event prediction with deep neural networks, in which a loss function that scales well to large data sets and enables the fitting of both proportional and non-proportional extensions of the Cox model. Discrete-time survival models have also gained great attention in the deep learning community due to their flexibility in handling censored data effectively and integrating seamlessly with neural network architectures. For the time-to-event data, it can be shown that by treating the time interval as an input variable in a standard feed-forward network with logistic activation and entropy error function, it is possible to estimate smoothed discrete hazards as conditional probabilities of failure.[9,10] Kvamme and Borgan showed that the discrete-time survival data can be parameterized based on the likelihood of a probability mass function with neural networks [11]. These discrete-time models offer various strategies for tackling survival analysis within the deep learning framework, each with its own strengths and considerations.

However, in clinical research, few studies intend to integrate diverse data sources, such as raw medical images, radiomic data, and clinical information, to develop accurate prognostic models. Traditional statistical methods often struggle with the high dimensionality and complex interactions inherent in these combined datasets. The deep learning methods mentioned earlier typically handle only single-source input data. Additionally, the time-to-event nature of survival data in cancer research adds complexity, requiring specialized analytical approaches for multi-source data prediction.

To tackle the aforementioned limitations in this study, inspired by the wide-and-deep learning model proposed by Cheng et al [12] for recommendation systems, we propose a novel method for survival prediction with multi-source data, which synergizes the strengths of discrete-time survival models and wide-and-deep learning architectures. Our approach, termed "discrete-time survival wide-and-deep learning," is designed to leverage both structured clinical and radiomic features and raw CT images to predict survival outcomes in lung cancer patients. This model aims to capture the complex interactions among clinical variables, imaging biomarkers, and raw images while accounting for the censored nature of survival outcomes. On the one hand, the wide component of our model incorporates clinical variables and engineered radiomic features, allowing for the capture of important linear and cross-product relationships. On the other hand, the deep component processes the raw CT images, extracting high-level semantic information and patterns that predefined radiomic features may not capture. By combining these components, we aim to create a comprehensive prognostic model that fully utilizes the diverse available patient data. Our study utilizes multi-resource datasets of 388 non-small cell lung cancer (NSCLC) patients,[13] including pretreatment CT images, high-dimensional radiomic features that quantify tumor image intensity, shape, and texture, and the patient's clinical data. This real-case application is to develop a discrete-time survival wide-and-deep learning model

that effectively integrates clinical, radiomic, and raw imaging data for survival prediction in lung cancer patients.

Specifically, the remaining paper is organized as follows. In Section 2, we review the discrete-time survival neural networks and then detail our proposed discrete-time survival wide-and-deep learning methodology. In Section 3, we present the real data analysis results for the lung cancer study. In Section 4, we conclude with a discussion and future directions of this work. Through this research, we aim to contribute to the growing body of knowledge at the intersection of artificial intelligence, radiomics, and oncology, with the ultimate goal of improving outcomes for lung cancer patients through more accurate and personalized prognostic predictions.

Methods

Discrete-time Survival Deep Learning

Given n patients, let T_i denote the event time and C_i the censoring time of individual i, $i = 1, \dots, n$. T_i and C_i are assumed to be independent random variables taking discrete values in $\{1, \dots, L\}$. For right-censored data, the observation time is defined by $\tilde{T}_i = \min(T_i, C_i)$, that is, $\tilde{T}_i$ corresponds to the observed event time, if $T_i < C_i$, and to the censoring time otherwise. In the discrete context, the discrete event times $0 < t_1 < t_2 < \cdots < t_L$. The originally continuous survival times are grouped into $l = 1,2, \dots, L$ disjoint intervals $(0, t_1]$, $(t_1, t_2]$, …, $(t_{L-1}, t_L]$. Denote $A_l = (t_{l-1}, t_l]$, with $t_0 = 0$ and l_i the last observation interval for the ith patient. Under this framework, the discrete probability function $f_l = P(T \in A_l) = S(t_{l-1}) - S(t_l)$, where the survival function $S(t_l) = P(T > t_l)$. The discrete hazard rate, h_l, is defined as the conditional failure probability

$$h_l = P(T \in A_l | T > t_{l-1}) = \frac{f_l}{S(t_{l-1})}.$$

Introducing the event indicator δ_{il}, which is set to 1 in the interval A_l containing the event of interest for uncensored subjects, and 0 otherwise, for the ith patient, the total likelihood of the discrete-time survival data is,

$$L = \prod_{i=1}^{n} \prod_{l=1}^{l_i} h_{il}^{\delta_{il}} (1 - h_{il})^{1-\delta_{il}} = \prod_{l=1}^{L} \prod_{i \in R_l} h_{il}^{\delta_{il}} (1 - h_{il})^{1-\delta_{il}}, \qquad (1)$$

where R_l is the set of patients at risk in the lth interval of time. Note that the likelihood (1) is a product of Bernoulli likelihoods. Therefore, one can fit the discrete-time survival model by running a logistic regression on a set of pseudo observations generated as follows: Suppose patient i dies or is censored at a discrete time point l_i. In addition, one observes a vector of p explanatory variables $\boldsymbol{x}_i = (x_{i1}, \dots, x_{ip})^T$. We treat the event indicator δ_{il} as the response variable. To each of these indicators for patient i, we associate a copy of the covariate vector $\boldsymbol{x}_i$ to each time interval. The discrete hazard rates are modeled by a logistic regression having the p covariates and the time interval as a block factor, i.e., $\text{logit}\left(h_l(\boldsymbol{x}_i)\right) = \theta_l + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}$, where $\theta_l, \beta_1, \dots, \beta_p$ are the parameters to be estimated.

To apply a feed-forward neural network as an alternative approach to logistic regression, we can take the negative logarithm of the likelihood (1), obtaining

$$l = -\log(L) = -\sum_{i=1}^{n} \sum_{l=1}^{l_i} (\delta_{il} \log h_l + (1 - \delta_{il}) \log(1 - h_l)).$$

It is equivalent to the cross-entropy error function in a binary classification problem in a neural network. We model $h_l(\cdot)$ with a logistic (sigmoid) function

$$h_l(\boldsymbol{x}_i, a_l) = \frac{\exp(z(\boldsymbol{x}_i, a_l))}{1 + \exp(z(\boldsymbol{x}_i, a_l))},$$

where $z(\cdot)$ is weighted sum of inputs into the neuron, and a_l represents a transformation of the time interval l. It can be a simple numeric representation of time or a more complex function such as a polynomial or spline transformation of time. For a simple case, a_l could be the midpoint of the interval l. For a given interval l, the weighted input z is calculated as

$$z(\boldsymbol{x}_i, a_l) = \alpha_0 + \sum_{m=1}^{M} v_m \sigma\left(\alpha_m + u_m a_l + \sum_{j=1}^{p} w_{mj} x_{ij}\right),$$

where $\alpha_0,\ v_1, \ldots, v_M,\ , \alpha_1, \ldots, \alpha_M,\ , u_1, \ldots, u_M,\ w_{11}, \ldots, w_{1p},\ \ldots\ ,\ w_{M1}, \ldots, w_{Mp},$ are the parameters to be estimated, and $\sigma(\cdot) = \exp(\cdot)/(1 + \exp(\cdot))$ is the activation function.

The discrete-time survival model with neural networks offers several notable advantages.[10] Firstly, its scalability allows it to efficiently handle large datasets, making it suitable for big data applications in healthcare and other fields. Secondly, the model's flexibility is a key strength, accommodating time-varying covariates and capturing non-linear relationships between predictors and outcomes. This adaptability is particularly valuable in complex medical scenarios where risk factors may change over time. Thirdly, another significant benefit is the model's interpretability; it provides easily understandable hazard and survival curves, which is crucial for clinical applications where transparency in decision-making processes is essential. Finally, unlike Cox-based models, this approach does not rely on the proportional hazards assumption, allowing it to model more diverse and realistic survival scenarios. These combined advantages make the model a powerful tool for survival analysis in various domains, especially in medical research where complex, time-dependent data is common.

Wide-and-Deep Survival Deep Learning

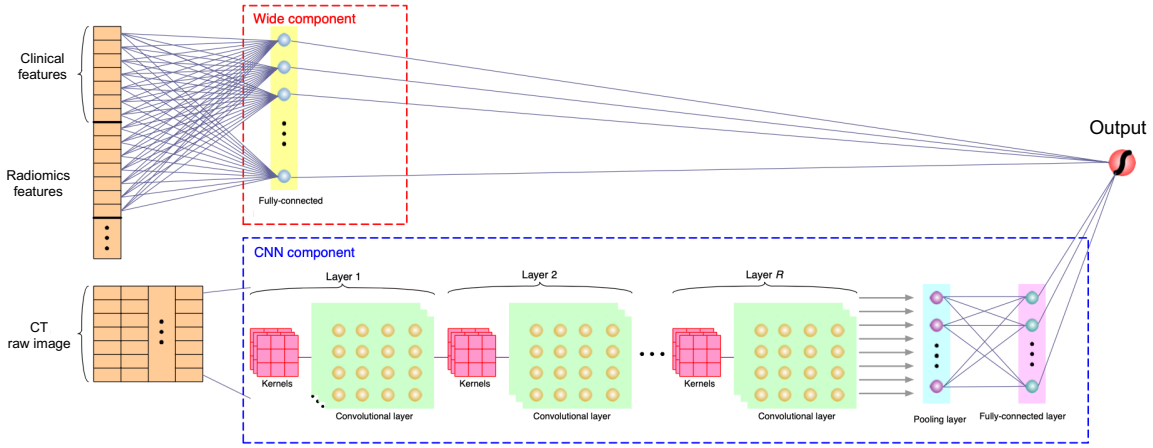

Figure 1. Discrete-time Survival wide-and-deep learning modeling framework for multi-resource data.

The wide-and-deep learning method was proposed by Cheng et al.[12], which integrates the strengths of linear models and deep neural networks into a cohesive framework for recommendation systems. It addressed a critical challenge in recommendation systems by balancing memorization and

generalization and demonstrated significant real-world impact through implementation on Google Play. The model combines a wide component, which focuses on memorization through cross-product feature transformations of sparse inputs, with a deep component that generalizes via low-dimensional dense embeddings. These components are jointly trained to optimize a common loss function, facilitating the model's ability to both exploit known interactions and discover novel patterns.

In the field of medical research, the construction of predictive models frequently involves the integration of multi-source data. This is particularly evident in our lung cancer research study, where we utilize not only clinical and demographic variables and engineered radiomic features but also the raw CT images for each patient. Inspired by the wide-and-deep learning model proposed by Cheng et al.[12] for recommendation systems, we introduce a novel method for survival prediction that leverages multi-source data, synergizing the strengths of discrete-time survival models with wide-and-deep learning architectures. Our approach, termed "Discrete-Time Survival Wide-and-Deep Learning," is strategically designed to harness both structured clinical and radiomic features alongside raw CT images to enhance the accuracy of survival outcome predictions in lung cancer patients. This model endeavors to capture the intricate interactions among clinical variables, imaging biomarkers, and raw images while diligently accounting for the censored nature of survival data.

Figure 1 shows the proposed network structure. In the proposed method, the wide component efficiently captures crucial linear relationships and cross-product interactions among clinical variables and radiomic features. This allows for a robust recognition of key feature interactions that predict survival outcomes. Concurrently, the deep component employs advanced 3D convolutional neural networks (CNNs) to process the raw CT images. This layer is designed to extract high-level semantic information and complex patterns not discernible through predefined radiomic features, thus enhancing the model's generalization capabilities. Specifically, the final patient-level latent representation, following the notations in Cheng et al. and Zaurin and Mulinka [12,14], is given by

$$z(\boldsymbol{x}_i, \boldsymbol{P}_i) = \boldsymbol{w}_{wide}^T CONCAT\big(\boldsymbol{x}_i, \phi(\boldsymbol{x}_i)\big) + \boldsymbol{w}_{deep}^T MLP\big(CNN(\boldsymbol{P}_i)\big),$$

where $\boldsymbol{x}_i$ and $\boldsymbol{P}_i$ are the tabular clinical and radiomic features and raw CT image for the patient i, *CONCAT* denotes vector concatenation, $\phi(\cdot)$ is a user-defined cross-product transformation for clinical and radiomic features to capture specific interactions, *CNN* denotes the raw image feature extractor from a convolutional network, *MLP* is a user-defined multi-layer perceptron, and $\boldsymbol{w}_{wide}^T$ and $\boldsymbol{w}_{deep}^T$ are weights to be learned. By integrating these components, our model offers a comprehensive and sophisticated prognostic tool that optimally leverages the full spectrum of available patient data.

A Real Case Study of Lung Cancer

In our real case study, we utilize a subset of the radiomics study conducted by Aerts et al.[13] to demonstrate the prediction of lung cancer patient survival with discrete-time survival wide-and-deep learning using high-dimensional radiomics features and raw CT images. The dataset comprises 388 non-small cell lung cancer (NSCLC) patients, with 292 patients in the training set and 96 in the testing set. For each patient, we have access to pretreatment CT scans, 3D gross tumor volume delineations by radiation oncologists, clinical features (age, gender, stage, histology), and the primary outcome variable (time-to-death). The original CT images and clinical data are available for download from The Cancer Imaging Archive (TCIA) portal (https://www.cancerimagingarchive.net). The radiomic dataset was obtained from Benelli et al.[15], in which radiomic feature extraction was employed with PyRadiomics[16], Version 2.2.0, with default parameters. The radiomic dataset contains a comprehensive set of 797 features, where these features capture various aspects of tumor characteristics, including shape, intensity, and texture.

To perform a thorough analysis of the radiomic features, we utilized the recently developed R software package RadAR (Radiomics Analysis with R)[15]. RadAR offers a complete pipeline for processing

radiomic datasets, from data import to feature processing and visualization. This package, in conjunction with other advanced R packages, forms the basis of our conventional machine learning modeling analysis. For our wide-and-deep learning approach, we implemented a custom algorithm using Python. This novel method integrates the strengths of both wide linear models and deep neural networks, allowing us to capture both global interactions between clinical features and radiomics features, as well as local semantic information from raw CT images.

We employed a comprehensive training and validation strategy for our deep and wide network. The model was trained using the Adam optimizer for 100 epochs, incorporating weight decay to mitigate overfitting. To ensure robust performance estimation and account for data variability, we implemented a repeated cross-validation approach. Specifically, the training data was randomly partitioned into two subsets: 80% for model training and 20% for validation. This process was repeated 10 times, each with a different random split, to obtain a more reliable estimate of the model's performance and stability. Hyperparameter tuning was conducted through an extensive grid search methodology. We evaluated the model's performance on the validation set across a range of hyperparameter combinations, including learning rate, batch size, network architecture (number and size of hidden layers), and regularization strength. The optimal hyperparameter configuration was selected based on the average performance across the 10 validation splits, ensuring that our final model was not overly sensitive to a particular data partition.

As a point of comparison, we implemented a random survival forest model[17] (using the R randomForestSRC package) as our baseline. This model was trained on the combined tabular clinical and radiomic data, providing a robust benchmark representative of traditional machine learning approaches in survival analysis. The random survival forest is known for its ability to capture non-linear relationships and interactions between features, making it a strong competitor for our deep learning approach.

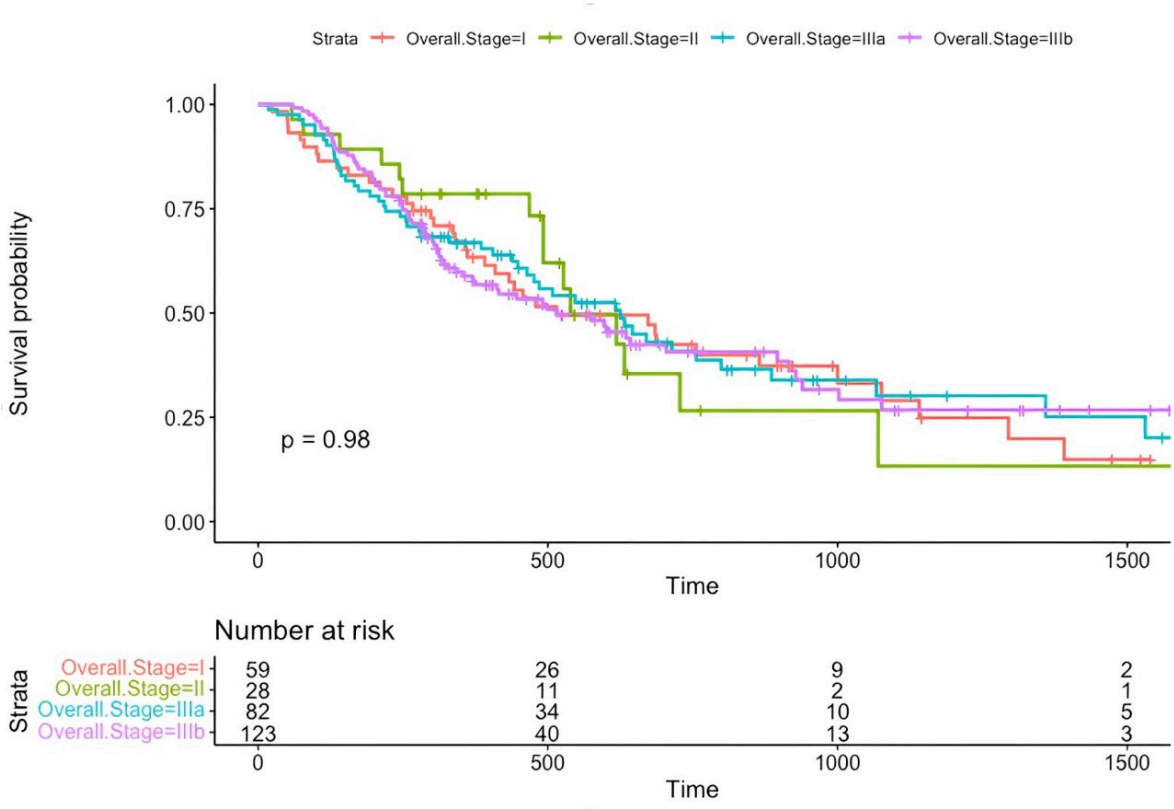

Figure 2. Kaplan-Meier curves for the time-to-death stratified by the cancer stage.

To assess the performance of our method, we utilized the integrated Brier score (IBS), a widely recognized metric in survival analysis. The IBS quantifies the model's predictive accuracy over the entire follow-up period, providing a comprehensive measure of performance. This score ranges from 0 to 1, with values closer to 0 indicating superior predictive accuracy. In addition to the IBS, we also evaluated the time-dependent Brier score curve. This complementary measure assesses the accuracy of survival predictions at specific time points, offering a more granular view of the model's performance over time. The time-dependent Brier score, also ranging from 0 to 1, allows us to identify periods where the model excels or potentially needs improvement. By employing both the integrated and time-dependent Brier scores, we gain a comprehensive understanding of our model's predictive capabilities.

Figure 2 illustrates the Kaplan-Meier survival curves stratified by cancer stages for our study's training cohort. Notably, the log-rank test comparing the cancer stage groups yielded a p-value of 0.98, indicating no statistically significant difference in survival outcomes among the different cancer stages. This unexpected result suggests that in our cohort, cancer stage alone may not be a reliable predictor of survival, underscoring the need for more sophisticated prognostic models.

We implemented a random survival forests model as our baseline, configuring it with 1,000 trees and a minimum terminal node size of 5. On the training set, this model achieved an integrated Brier score (IBS) of 0.19, indicating reasonably good predictive performance. However, when applied to the testing set, the IBS increased substantially to 0.36. This marked increase suggests potential overfitting of the model to the training data, highlighting the challenges in developing generalizable survival prediction models. Figure 3 presents the variable importance plot for the random survival forests model. The analysis identified 32 variables that positively contributed to the model's predictive power. Interestingly, among these influential variables, age emerged as the sole important clinical contributor. The remaining significant predictors were predominantly radiomic features, emphasizing the potential value of quantitative imaging analysis in survival prediction.

We implemented our proposed wide-and-deep learning model which integrates radiomic and clinical data with raw CT image data, leveraging the strengths of both structured and unstructured information. The results were promising, with the wide-and-deep model achieving an IBS of 0.17 on the testing set, a substantial improvement over the random forest model's performance. Figure 4 displays the time-dependent Brier score curve for the testing set, offering insights into the model's accuracy at clinically relevant time points. This curve demonstrates the superior performance of our wide-and-deep learning approach across various time horizons, further validating its effectiveness.

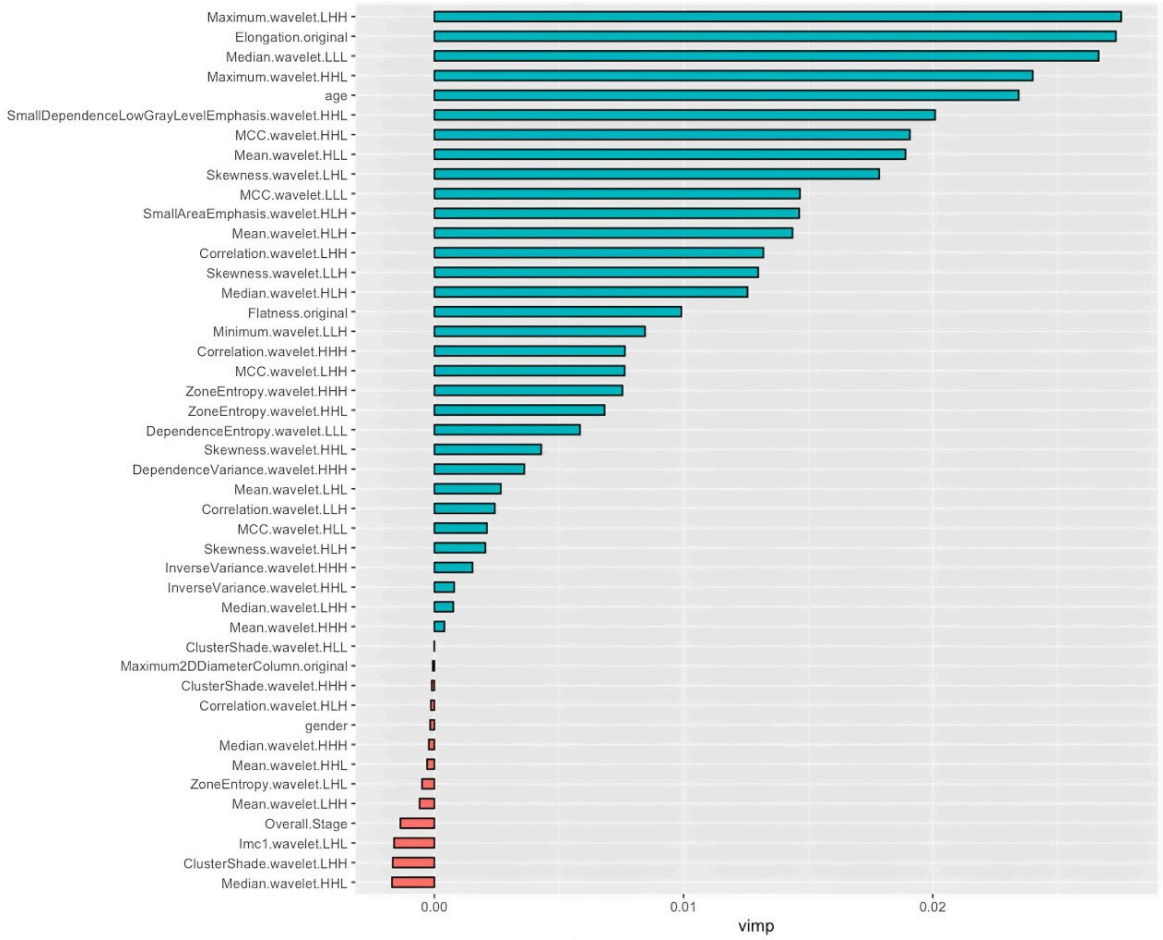

Figure 3. Variable importance plot based on the survival random forest model.

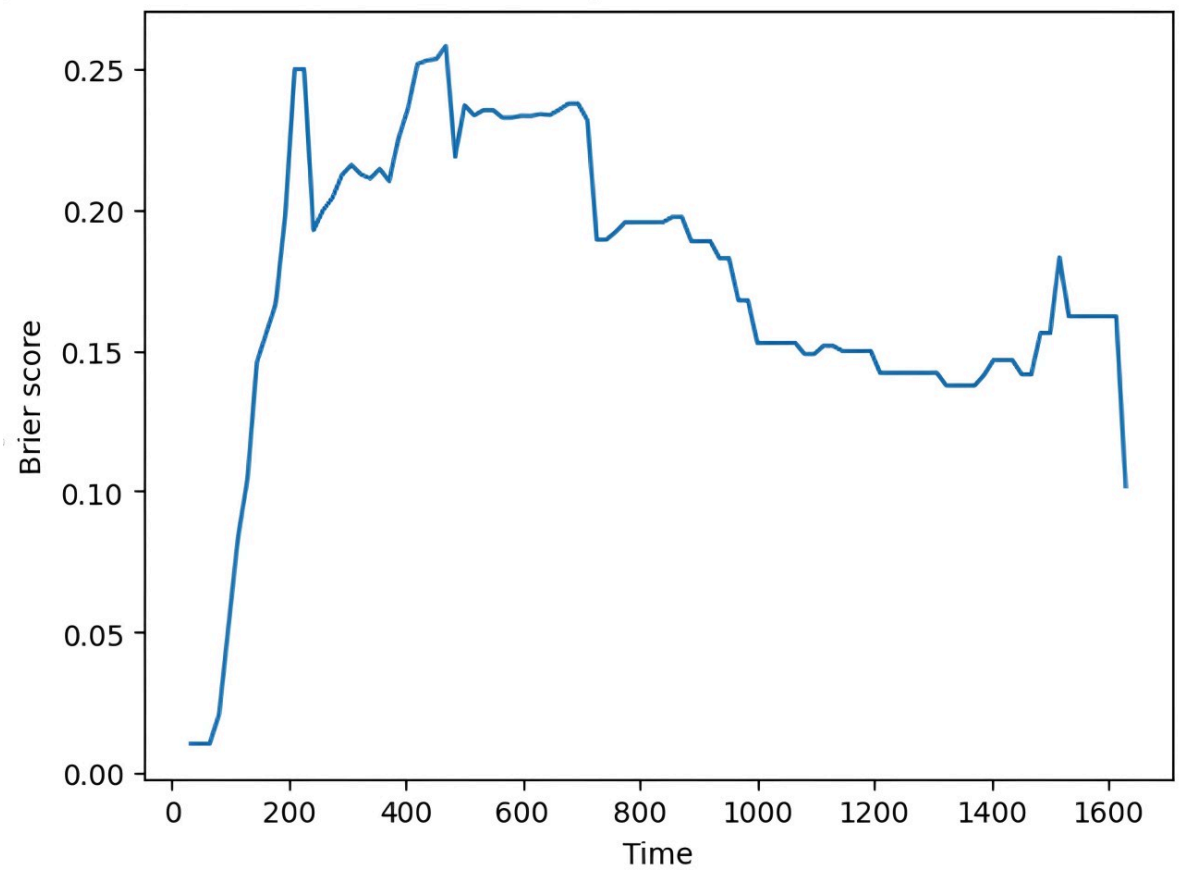

Figure 4. Time-dependent Brier score curve for the testing set using the wide-and-deep learning model. The IBS for the testing set is 0.17.

Discussion

In this study, we proposed a novel discrete-time survival wide-and-deep learning framework for survival prediction that effectively leverages multi-source data. Our approach innovatively combines the strengths of traditional survival analysis with advanced deep learning techniques, addressing the complex challenges presented by high-dimensional, multi-modal medical data.

The wide component of our model focuses on efficiently learning frequent and specific feature interactions, particularly from clinical and engineered radiomic features. This allows for the capture of important linear and cross-product relationships that are crucial for accurate survival prediction. Concurrently, the deep component utilizes advanced 3D CNNs with embeddings to foster generalization. This enables the model to learn and predict unseen feature combinations through low-dimensional dense embeddings, extracting high-level semantic information and patterns from raw CT images that predefined radiomic features might not capture.

A key innovation of our approach is the utilization of a discrete-time survival likelihood strategy. This method effectively addresses the challenge posed by the censored nature of survival outcomes, ensuring that our model remains sensitive to the temporal aspects of patient survival data. By doing so, we maintain the integrity of time-to-event information while harnessing the power of deep learning architectures.

Our real case study on lung cancer patients provides compelling evidence that the proposed wide-and-deep learning approach significantly outperforms traditional methods in predicting patient survival.

The model's superior performance is demonstrated by its lower integrated Brier score and consistently better time-dependent Brier scores compared to the random survival forest baseline. This improvement in predictive accuracy is achieved by effectively capturing and integrating diverse factors influencing patient outcomes, including subtle imaging features that may be overlooked by conventional analyses.

The success of our model underscores the potential of advanced machine learning techniques in enhancing personalized cancer care and treatment planning. By providing more accurate and comprehensive prognostic information, our approach could aid clinicians in making more informed decisions about treatment strategies and patient management. Furthermore, the model's ability to integrate multi-modal data – clinical, radiomic, and raw imaging – offers a more holistic view of patient health, potentially uncovering novel prognostic factors and biomarkers.

However, it is important to acknowledge the limitations of our study. The unexpected lack of prognostic value in cancer staging in our cohort highlights the potential for dataset-specific anomalies and emphasizes the need for extensive validation across diverse patient populations. Additionally, while our model shows improved performance, the interpretability of deep learning models remains a challenge, which is crucial for clinical adoption.

Future work should focus on external validation of our model across larger and more diverse patient cohorts to ensure generalizability. Efforts to enhance model interpretability, perhaps through techniques like attention mechanisms or feature importance analysis for the deep component, would be valuable for clinical translation. Additionally, investigating the model's performance in predicting other outcomes, such as treatment response or recurrence-free survival, could further demonstrate its utility in oncology.

Conclusion

In conclusion, our discrete-time survival wide-and-deep learning framework represents a significant advancement in survival analysis for complex medical data. By effectively integrating multi-source data and leveraging both the memorization capabilities of wide models and the generalization power of deep learning, our approach offers a promising tool for enhancing prognostic accuracy in lung cancer and potentially other areas of oncology. This work contributes to the growing intersection of artificial intelligence, radiomics, and personalized medicine, paving the way for more accurate, individualized patient care in the era of precision oncology.

Acknowledgments

The author appreciates Dr. Penglei Gao's guidance, which made the completion of this project possible. The author also appreciates the internship opportunities provided by the Cleveland Clinic Lerner Research Institute.

References

1. Jonas DE, Reuland DS, Reddy SM, et al. Screening for Lung Cancer With Low-Dose Computed Tomography: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA.* 2021;325(10):971-987.
2. National Lung Screening Trial Research Team, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med.* 2011;365(5):395-409.
3. Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med.* 2019;25(6):954-961.
4. Bogie K, Wang X, Fei B, Sun J. New technique for real-time interface pressure analysis: getting more out of large image data sets. *J Rehabil Res Dev.* 2008;45(4):523-535.

5. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology.* 2016;278(2):563-577.
6. Wang X, Pennello G, deSouza NM, et al. Multiparametric Data-driven Imaging Markers: Guidelines for Development, Application and Reporting of Model Outputs in Radiomics. *Acad Radiol.* 2023;30(2):215-229.
7. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC medical research methodology.* 2018;18(1):24.
8. Kvamme H, Borgan Ø, Scheel I. Time-to-event prediction with neural networks and Cox regression. *Journal of Machine Learning Research.* 2019;20(129):1-30.
9. Biganzoli E, Boracchi P, Mariani L, Marubini E. Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach. *Statistics in medicine.* 1998;17(10):1169-1186.
10. Gensheimer MF, Narasimhan B. A scalable discrete-time survival model for neural networks. *PeerJ.* 2019;7:e6257.
11. Kvamme H, Borgan O. Continuous and discrete-time survival prediction with neural networks. *Lifetime Data Anal.* 2021;27(4):710-736.
12. Cheng H-T, Koc L, Harmsen J, et al. Wide & deep learning for recommender systems. *Proceedings of the 1st workshop on deep learning for recommender systems.* 2016:7-10.
13. Aerts H, Velazquez ER, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature communications.* 2014;5(1):1-9.
14. Zaurin JR, Mulinka P. pytorch-widedeep: A flexible package for multimodal deep learning. *Journal of Open Source Software.* 2023;8(86):5027.
15. Benelli M, Barucci A, Zoppetti N, et al. Comprehensive analysis of radiomic datasets by RadAR. *Cancer Research.* 2020;80(15):3170-3174.
16. Van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer research.* 2017;77(21):e104-e107.
17. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *The annals of applied statistics.* 2008;2(3):841-860.

Author

Steven Wang is a senior at Solon High School in Solon, OH, United States. Since 2023, he has actively engaged as a research student in several labs at the Cleveland Clinic Lerner Research Institute in Cleveland, Ohio. Steven is particularly passionate about the intersection of medicine and artificial intelligence, with a keen interest in leveraging machine learning and deep learning algorithms to improve clinical disease prognostic predictions. His dedication to research is evident through his contributions to various projects aimed at enhancing patient outcome predictions through novel computational techniques.