Survival Analysis by Noninvasive Imaging: A Quantitative Radiomic Approach

By AKASH KUMAR GUPTA

WORD COUNT 3974
TIME SUBMITTED 22-JUL-2022 02:27PM
PAPER ID 88696646

Survival Analysis by Noninvasive Imaging: A Quantitative Radiomic Approach

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Technology
in
Computer Science

by

Akash Kumar Gupta
CS2003

under the guidance of

Dr. Dipti Prasad Mukharjee
Professor
Electronics and Communication Sciences Unit

Indian Statistical Institute
Kolkata-700108, India

July 2022

To my family and my guide

CERTIFICATE

This is to certify that the dissertation entitled **"Survival Analysis by Noninvasive Imaging: A Quantitative Radiomic Approach"** submitted by **Akash Kumar Gupta** to Indian Statistical Institute, Kolkata, in partial fulfillment for the award of the degree of **Master of Technology in Computer Science** is a bonafide record of work carried out by him under my supervision and guidance. The dissertation has fulfilled all the requirements as per the regulations of this institute and, in my opinion, has reached the standard needed for submission.

21/07/2022

Dr. Dipti Prasad Mukherjee
Professor(HAG),
Electronics and Communication Sciences Unit,
Indian Statistical Institute,
Kolkata-700108, INDIA.

ABSTRACT

Strong phenotypic variations in lung tumors of human can be seen non-invasively via medical imaging. These tumor phenotype can be quantified by extracting large number image features and is referred as radiomics. It has already been shown that the radiomic signature of a tumor has significant association with some of its clinical parameters [2]. These radiomic characteristics may be associated to the survival of the patient, which is studied here.

A large number of statistical techniques for the survival analysis have been created over the years. Cox proportional hazard model [6] has been used to explore the possible association of radiomic signature and clinical parameters with the survial function. Radiomic signatures shows significant association with the survival as compared to the clinical parameters [2].

Looking into the recent development of machine learning and neural networks, improved method to estimate the survival function has been developed based on discrete-time survival likelihood using neural networks [12]. The survival method under consideration parameterize the discrete-time hazard rate depending on the likelihood for right-censored survival data as well [13] with neural network. This neural network predicts the survival curve using hand calculated radiomic features as the covariates. As hand calculated features may not capture all aspects of the tumor, a model based on convolution neural network is proposed to extract radiomic features from tumor volume to predict the survival function. Radiomic signature extracted from the CNN shows marginal improvement over the hand calculated radiomic signature in survival analysis and gives better insight to the association of the tumor radiomics with survival.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Computed Tomography Scan

In clinical practise, computed tomography (CT) is frequently used to determine lung cancer treatment. It is a key diagnostic tool for an oncologist that allowed for risk assessment. For patients with lungs cancer, CT scan offers variety of data, that influence the treatment procedure. However, a qualitative interpretation of CT scan is still restricted to what an expert radiologist may see. This motivates to find a need for a strong interpretation of CT scan.

CT scans of patients with lung cancer and others, have been shown that it include more information than what is visible to radiologists using quantitative image analysis [8]. Recent developments in machine learning, particularly convolutional neural networks, have produced a class of potent models that show promise for improving medical decision-making and achieving accurate diagnosis [17]. Using CNN-based models on imaging data can reveal clinically significant prognostic patterns that were previously unknown or invisible to the eye of an expert.

1.2 Survival Analysis

A statistical technique called survival analysis is used to determine how long it will likely be before a given event occurs [15]. These events could be something like a death in the medical area. The survival times are typically expressed in

terms of days, weeks, months or years. The years till a person's death, for instance, may be considered the survival time if the event of interest is death. The survival function and the hazard function are the other two key functions that make up a survival analysis. The probability that a person has "survived" past time t is represented by the survival function, $S(t) = P(T > t)$. A measurement of risk at time t is the hazard function, $h(t)$. A higher hazard ratio indicates a higher danger of dying.

1.3 Problem Statement and Contribution

Tumors are heterogeneous, hence the biopsy or invasive extraction of tissues usually from a very small area are not enough to characterize the tumor. Hence, as the non invasive techniques like CT Scan, capture the comprehensive information of the whole tumor,it is used to track the growth of the illness or how it responds to treatment.

In this study, we have focused on survival analysis of confirmed non small cell lungs cancer (NSCLC) patients using the information extracted non invasively from their CT scan. We developed a neural network and a convolution neural network to analyse CT images and estimate the survival function for that patient. The most popular method for survival analysis in a wide range of domains is the Cox Proportional Hazards (CPH) model [6]. Due to its simplicity of use and speedy computation, the Cox regression model has been widely accepted by the scientific community. But it has some flaws. For example, it is a poor model for high dimensions [3]. Also, it assumes that the hazard ratio for two patients is constant with respect to time. It has inability to model the

non linearities. A number of studies have consistently showed that ML-based methods may predict patient survival at least as well as traditional CPH analysis [11] [16]. We used an NN based model and developed a CNN based model to estimate the survival curve by analysing the CT scan of NSCLC patients.

1.4 Organization

In chapter 2, we have discussed the previous works done in the field of survival analysis. In chapter 3, we have explained the proposed methodology with an elaborate explanation. We explain the dataset we used for the study. We explain the architectures of NN and CNN model used along with the loss function used to optimize those models. Finally in chapter 4, we discussed the results and conclude our study.

CHAPTER 2

RELATED WORK

The survival analysis is frequently used to forecast a patient's prognosis. Numerous research have been carried out using the Kaplan-Meier [10] survival estimator and the Cox hazard model [6].

One of the most often used approaches for estimating survival is the Kaplan-Meier estimator [10]. It can't be used to estimate an individual's survival time because it does not incorporates any of the patient's variables (e.g. clinical parameters). It predicts the survival distribution function only from survival data.

The Cox hazard model calculates an individual's hazard ratio and assesses the impact of patient's covariates on the survival models. Different modified Cox models have been proposed over the years, for instance the Cox Boost algorithm [4], Lasso Cox model [20] and many others. Some new model based on deep learning methods that uses cox regression model loss function have been developed like DeepSurv [11] and DeepHit [14] and they outperforms the conventional methods. However, a specific family distribution for the hazard function has to be selected to get the survival function for a patient.

CHAPTER 3

METHODOLOGY

3.1 Overview

Using the non invasive imaging technique, we want to estimate the survival function of the patient diagnosed with non small cell lung cancer (NSCLC). We used hand calculated features as well features extracted using convolution neural network. A neural network model is trained using hand calculated features. A separate Convolution neural network is also trained. Survial estimates for both of these models are compared with that of classical Cox proportional hazard regression. Fig.3.1 shows the basic idea of the work flow.

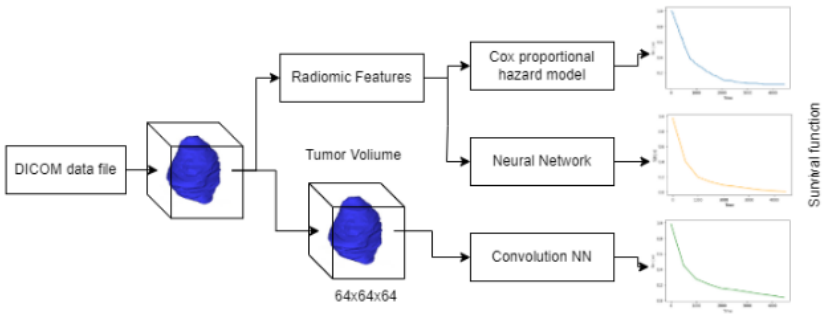


Figure 3.1: Analysis workflow: The radiomics features extracted from the tumor volume is used to train CPH and NN model the survial probability with respect to the time. Tumor volume is used in CNN model to get the same output.

3.2 Data

Lung1 data set: The Lung1 dataset includes clinical information and computed tomography (CT) scans from 422 non-small cell lung cancer (NSCLC) patients

who underwent radiotherapy [1] at MAASTRO Clinic in Netherlands. CT images, manual delineations of tumor, clinical information, and survival data were available for these patients. However, as we are estimating the survival curves using non invasive techniques, we only use the CT scans and survival data of the patients. The resolution for CT scan was 512×512. Each patient has different number of CT scan slices ranging from 90 to 150. Survival data includes number of days of survival after diagnosis and their death status (death status = 0 if still alive and is called right-censored). After the pre-processing, data for 414 patients [1] were used in this study.

3.2.1 Data Pre-processing

The CT scans were available in DICOM format. DICOM Segmentation (SEG) files included in this data contain the manually delineation of different anatomy (i.e., hearts, lungs, neoplasm and esophagus) by radiation oncologist. The 'neoplasm' label from DICOM-SEG was the tumor in which we are interested and was used to create the 3d mask for the tumor volume. Using this mask and the CT scan of corresponding patients, 3d volume was generated.

Feature Extraction: For the analysis of survival using cox-regression and neural networks, 808 features quantifying tumor shape, intensity and textures were extracted using pyradiomics [19].Features including first order statistics (e.g.,mean, deviation, energy, entropy, etc.), shape based features (e.g., volume, surface area, compactness, etc.) and for textures based on gray level cooccurrence matrix, gray level run length matrix, gray level side zone matrix and gray level dependence matrix were extracted from the original volume. The first or-

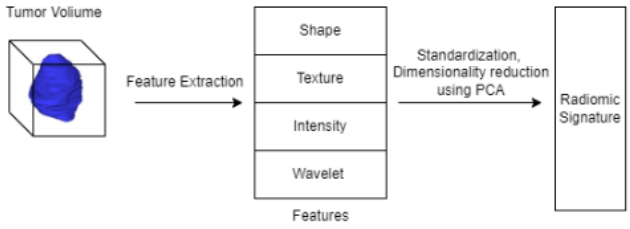


Figure 3.2: Building hand crafted radiomic signature

der statistics and texture based features were also extracted from the 8 volumes decomposed using wavelet (high pass and low low pass in each of the three axis). These features were standardized for further processing. To eliminate the redundancy in the feature group, dimensions were reduced using Principal component analysis (PCA). 64 eigen vectors were selected as they collectively explained the 98.16% of the total variance in the data set. We refer the features projected on these new 64 dimensions as hand calculated radiomic signatures as shown is Fig.3.2. We used the radiomic signature in the survival analysis using Cox regression and neural network.

Preparation of data for convolution neural network: The size of the tumor and slices of the CT scan for each patient is different. Consequently, the dimensions of the 3d tumor volume is different for each patients. For the CNN, we needed every tumor volume be made to same dimension. To do that, we cropped up the boundary box of the tumor. Then we zero padded in every direction symmetrically such that the dimension of the 3d volume is 512×512×512. Now, the 3d volume is resized to $64 \times 64 \times 64$. This 3d tumor volume is used in survival analysis of patient using CNN.

3.3 Neural Network Architecture

The framework for NN architecture is show in the Fig.3.3. This network takes the radiomic signature (x) as input. The input layer has 64 neurons. There are 4 hidden layers in the network . Each is Leaky-ReLU activated with batch normalisation and with dropout probability of 0.1. This network was trained with a batch-size of 32. The output layer consists of 10 neurons each representing a discrete time point (t_i). Each neuron of output layer gives the hazard rate corresponding to that discrete time. As the hazard function ($h(t_i)$) as defined in section 3.5.1, is a probability, its value should lie between 0 and 1. Applying the logistic function (sigmoid function) to the output ($\phi(x)$) of neural network will do this. Survival probabilities corresponding to the discrete time can be then calculated as using equation 3.3. The model was implemented using pytorch and the snapshot of the implemented is shown in Fig.3.4. The 'Layer' column describes the every layer of the NN model, 'Param #' describes the number of parameter required for the corresponding layer.

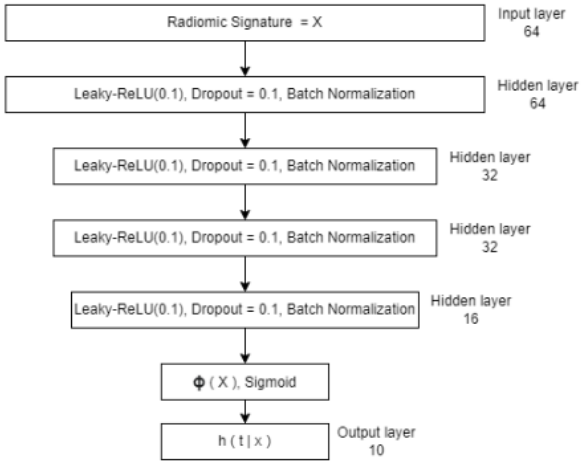
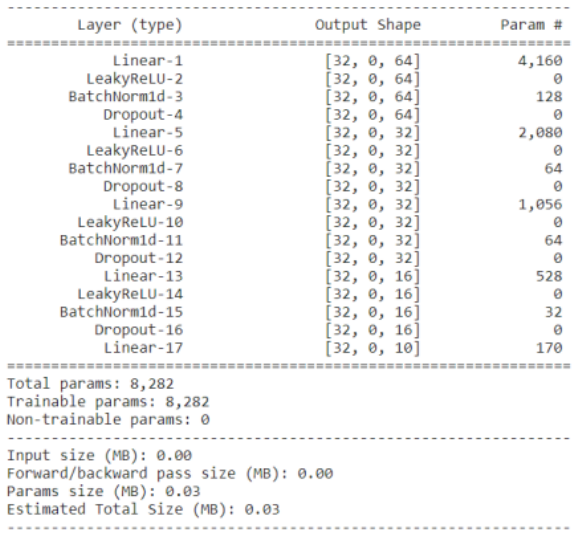


Figure 3.3: NN architecture model used for survival analysis



Layer (type)	Output Shape	Param #
Linear-1	[32, 0, 64]	4,160
LeakyReLU-2	[32, 0, 64]	0
BatchNorm1d-3	[32, 0, 64]	128
Dropout-4	[32, 0, 64]	0
Linear-5	[32, 0, 32]	2,080
LeakyReLU-6	[32, 0, 32]	0
BatchNorm1d-7	[32, 0, 32]	64
Dropout-8	[32, 0, 32]	0
Linear-9	[32, 0, 32]	1,056
LeakyReLU-10	[32, 0, 32]	0
BatchNorm1d-11	[32, 0, 32]	64
Dropout-12	[32, 0, 32]	0
Linear-13	[32, 0, 16]	528
LeakyReLU-14	[32, 0, 16]	0
BatchNorm1d-15	[32, 0, 16]	32
Dropout-16	[32, 0, 16]	0
Linear-17	[32, 0, 10]	170

Total params: 8,282
Trainable params: 8,282
Non-trainable params: 0

Input size (MB): 0.00
Forward/backward pass size (MB): 0.00
Params size (MB): 0.03
Estimated Total Size (MB): 0.03

Figure 3.4: Snapshot of the summary of the NN model implemented in pytorch

3.4 Convolution Neural Network Architecture

The proposed framework for CNN architecture is shown in Fig.3.5. This network take the 3d volume of the tumor as input. The size of the input is $64 \times 64 \times 64$. The first hidden layer produces 16 output features maps using kernel of size $3 \times 3 \times 3$ using 3d convolution. The same procedure is repeated for next two layers. The last feature map is then flattened and is fully connected to 2 more hidden layers with 512 and 64 nodes respectively. The output layer consists of 10 neurons each representing a discrete time point. Again sigmoid function is used to get the values between 0 and 1 for the hazard function. Survival probabilities corresponding to the discrete time can be then calculated using equation 3.3. This model was implemented using the pytorch Fig.3.6 shows the snapshop of the architecture of the model. Here again 'Layer' column shows all the layers in the CNN model. In 'Output Shape', first number is the batch

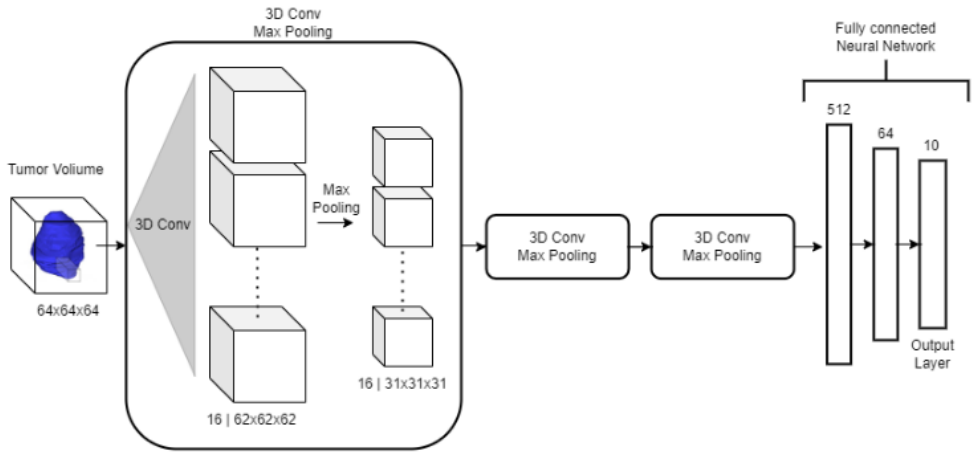


Figure 3.5: CNN architecture model used for survival analysis

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv3d-1       [1, 16, 62, 62, 62]             448
         MaxPool3d-2       [1, 16, 31, 31, 31]               0
            Conv3d-3       [1, 16, 29, 29, 29]           6,928
         MaxPool3d-4       [1, 16, 14, 14, 14]               0
            Conv3d-5       [1, 16, 12, 12, 12]           6,928
         MaxPool3d-6          [1, 16, 6, 6, 6]               0
            Linear-7                [1, 512]       1,769,984
         LeakyReLU-8                [1, 512]               0
           Dropout-9                [1, 512]               0
           Linear-10                 [1, 64]          32,832
        LeakyReLU-11                 [1, 64]               0
          Dropout-12                 [1, 64]               0
           Linear-13                 [1, 10]             650
        LeakyReLU-14                 [1, 10]               0
          Dropout-15                 [1, 10]               0
================================================================
Total params: 1,817,770
Trainable params: 1,817,770
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 1.00
Forward/backward pass size (MB): 36.29
Params size (MB): 6.93
Estimated Total Size (MB): 44.23
----------------------------------------------------------------
```

Figure 3.6: Snapshot of the summary of the CNN model implemented in pytorch

size, and the second number is the number of kernels in that layer. The last three numbers are the dimensions of the feature map. For linear layer, first number is the batch size and the second number is the number of nodes. 'Param #' column contains the number of parameters required for that layer.

3.5 Loss Function

We begin by providing a quick overview of key terminologies used in the field of survival analysis. This is followed by the discussion on the loss function for right-censored survival data and discrete time model as discussed by Tutz (2016) [18] and Lee (2018) [14]. The implementation of this loss function was available in *pycox* library.

3.5.1 Background

Let us assume that the time has discrete value $t_j \in \mathcal{T}$ where $\mathcal{T} = \{t_1, t_2, ..., t_m\}$. The time of event (in our case death of the patient) is denoted as $T \in \mathcal{T}$. Let the pmf of such event is

$$f(t_j) = P(T = t_j) \tag{3.1}$$

The survival function is defined as the probability of a patient to survive past a certain time t_j. It can be written as

$$S(t_j) = P(T > t_j) \tag{3.2}$$

The hazard function is defined as the probability of death of a patient at time t_j given that the patient survived till t_{j-1}. It can be denoted as

$$h(t_j) = P(T = t_j | T > t_{j-1}) = \frac{f(t_j)}{S(t_{j-1})} = \frac{S(t_{j-1}) - S(t_j)}{S(t_{j-1})} \tag{3.3}$$

Keeping in mind the above equation, it implies that the expression for the survival function can be rewritten as

$$S(t_j) = \prod_{k=1}^{j} (1 - h(t_k)) \tag{3.4}$$

Kvamme and Borgan (see [12] for more information) derived the loss function using the mean negative log-likelihood which is given by

$$loss = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k(t_i')}(y_{ij}\log(h(t_j|x_i)) + (1 - y_{ij})\log(1 - h(t_j|x_i))) \tag{3.5}$$

Here,for a patient i, x_i is the radiomic signature. $k(t') \in \{0, 1, ..., m\}$ defines the discrete time index of time t' (i.e., $t' = t_{k(t')}$). y_{ij} corresponds to the event status at time t_j (i.e. death status of the i^{th} patient) We identify this loss function as the negative log likelihood for Bernoulli data, often known as the binary cross-entropy, a significant finding first made by Brown [5].

We minimise this loss function in out neural network for the purpose of estimating hazard function. Survival function then can be obtained using equation 3.4. Considering $\phi(x) \in \mathcal{R}^m$) be the output of the neural network corresponding to the m discrete time t. As hazards are probabilities, and require $h(t|x) \in [0, 1]$, we use logistic function to the output layer ($\phi(x)$) of the neural network to make it possible [7].

CHAPTER 4

RESULTS AND CONCLUSIONS

For training and validation of all the models, the data set was divided in training set of 265 patients and validation set of 66 patients. A test set 83 patients was used as an independent dataset for the testing and comparing the models. All the models were train using the same set of training, validation and testing datasets. All the models were trained in Kaggle Notebook, which is a cloud computational environment. The training and validation loss with respect to epoch for NN and CNN model is shown in Fig.4.1.

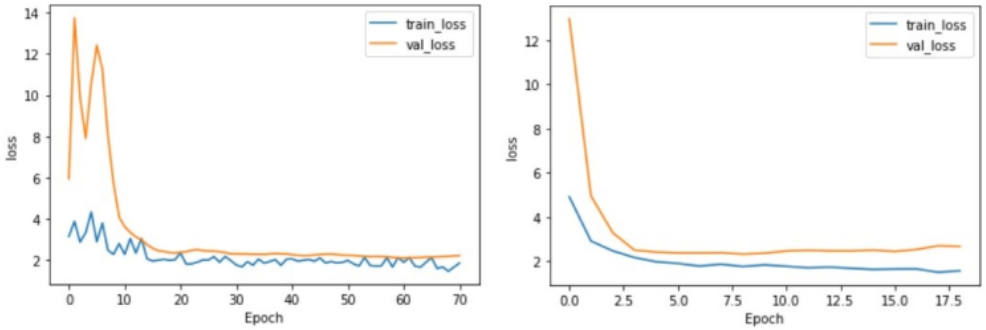


Figure 4.1: Training and validation loss vs epoch for NN (left) and CNN (right).

4.1 Comparing the Survival Curves

The survival probability with respect to the time is estimated for two example patients and are shown in Fig.4.2. In the left graph of Fig.4.2, the survival curves of a patient estimated by the CNN and NN models are very close as compared to that of Cox regression, On the other hand, the right graph of Fig.4.2, we can see the curve estimated by the CNN model is quite close to that of Cox regression

as compared to the curve estimated by NN model.

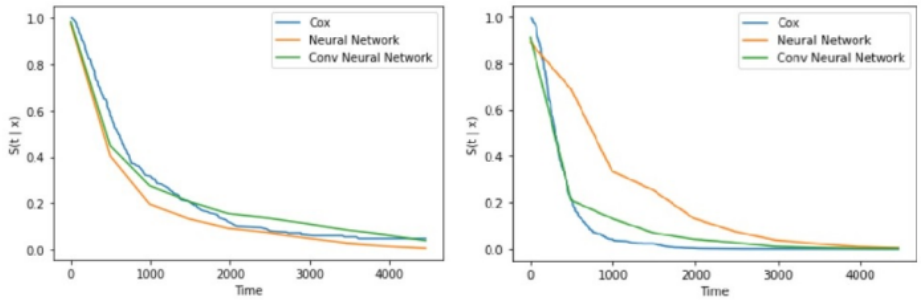


Figure 4.2: Survival function of two example patient estimated by NN and CNN compared against Cox regression. The time axis are in days.

4.2 Evaluating the Model

The model outputs the survival function. The area under the survival function is inversely proportional to the risk score (higher the area under the curve, lower the risk score). This indicates that the model can predict the sequence of event based on the area (i.e. sequence of the patient more likely to pass away). Therefore, We employ Harrell's concordance index to evaluate our model [9].

***c*-Index:** The *c*-index evaluates the rank correlation between the estimated risk score from the models and the actual time point of the event (in our case, death of the patient). A *c*-index of 0.5 a random estimation of life expectancy, where as a *c*-index of 1 means that the model is able to perfectly predict the risk for the patient. The formulation of *c*-index is as follows:

$$c - index = \frac{\text{number of concordant pairs}}{\text{total number of possible pairs}} \tag{4.1}$$

The pair here is considered to be actual time of survival of the patient and the area under the survival function of that patient estimated by the model.

Model	Concordance Index
Cox regression	0.5863
Neural Network	0.6248
Convolution Neural Network	0.6316

Table 4.1: Evaluation using c-index of different models

4.3 Conclusions

Table 4.1 shows the c-index of all the three models. As we can observe, the neural network performs much better then classical cox proportional hazard regression with the hand calculated radiomic signature as input. The proposed model using CNN performs similar to the neural network. There was no significant difference the performance as the c-indices for both models as quite close. However, reducing the dimension of the tumor volume from $512 \times 512 \times 512$ to merely $64 \times 64 \times 64$ for the CNN, might result in loss of some important features. The hand crafted features were calculated from full resolution images, hence the NN did not have that problem. The dimension of the tumor was needed to be reduced in order to lower the training parameters of the CNN model. Due to flexibility of the NN and CNN model as compared to the Cox regression, we get a higher performance. In conclusion, we showed that the invisible information in CT scan, a non invasive imaging technique, can be quantified and are clinically important as well.

BIBLIOGRAPHY

[1] Data from nsclc-radiomics (lung1). https://wiki.cancerimagingarchive.net/x/FgL1. Accessed: 21st March 2022.

[2] Hugo J. W. L. Aerts, Emmanuel Rios Velazquez, Ralph T. H. Leijenaar, Chintan Parmar, Patrick Grossmann, Sara Carvalho, Johan Bussink, René Monshouwer, Benjamin Haibe-Kains, Derek Rietveld, Frank Hoebers, Michelle M. Rietbergen, C. René Leemans, Andre Dekker, John Quackenbush, Robert J. Gillies, and Philippe Lambin. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*, 5(1):4006, Jun 2014.

[3] Axel Benner, Manuela Zucknick, Thomas Hielscher, Carina Ittrich, and Ulrich Mansmann. High-dimensional cox models: The choice of penalty as part of the model building process. *Biometrical Journal*, 52(1):50–69, Feb 2010.

[4] Harald Binder and Martin Schumacher. Allowing for mandatory covariates in boosting estimation of sparse high-dimensional survival models. *BMC Bioinformatics*, 9(1), Jan 2008.

[5] Charles C. Brown. On the use of indicator variables for studying the time-dependence of parameters in a response-time model. *Biometrics*, 31(4):863–872, 1975.

[6] D. R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2):187–220, 1972.

[7] Michael F. Gensheimer and Balasubramanian Narasimhan. A scalable discrete-time survival model for neural networks. *PeerJ*, 7:e6257, Jan 2019.

[8] Robert J. Gillies, Paul E. Kinahan, and Hedvig Hricak. Radiomics: Images are more than pictures, they are data. *Radiology*, 278(2):563–577, Feb 2016.

[9] Jr Harrell, Frank E., Robert M. Califf, David B. Pryor, Kerry L. Lee, and Robert A. Rosati. Evaluating the Yield of Medical Tests. *JAMA*, 247(18):2543–2546, May 1982.

[10] E. L. Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481, 1958.

[11] Jared L. Katzman, Uri Shaham, Alexander Cloninger, Jonathan Bates, Tingting Jiang, and Yuval Kluger. DeepSurv: personalized treatment recommender system using a cox proportional hazards deep neural network. *BMC Medical Research Methodology*, 18(1), Feb 2018.

[12] Hvard Kvamme and Ørnulf Borgan. Continuous and discrete-time survival prediction with neural networks, 2019.

[13] Hvard Kvamme, Ørnulf Borgan, and Ida Scheel. Time-to-event prediction with neural networks and cox regression. 2019.

[14] Changhee Lee, William Zame, Jinsung Yoon, and Mihaela Van der Schaar. DeepHit: A deep learning approach to survival analysis with competing risks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), Apr 2018.

[15] Elisa T. Lee and Oscar T. Go. Survival analysis in public health research. *Annual Review of Public Health*, 18(1):105–134, 1997. PMID: 9143714.

[16] Arturo Moncada-Torres, Marissa C. van Maaren, Mathijs P. Hendriks, Sabine Siesling, and Gijs Geleijnse. Explainable machine learning can outperform cox regression predictions and provide insights in breast cancer survival. *Scientific Reports*, 11(1), Mar 2021.

[17] Hoo-Chang Shin, Holger R. Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M. Summers. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5):1285–1298, May 2016.

[18] Gerhard Tutz and Matthias Schmid. *Modeling Discrete Time-to-Event Data*. Springer International Publishing, 2016.

[19] Joost J.M. van Griethuysen, Andriy Fedorov, Chintan Parmar, Ahmed Hosny, Nicole Aucoin, Vivek Narayan, Regina G.H. Beets-Tan, Jean-Christophe Fillion-Robin, Steve Pieper, and Hugo J.W.L. Aerts. Computational radiomics system to decode the radiographic phenotype. *Cancer Research*, 77(21):e104–e107, Oct 2017.

[20] Lu Wang, Jincheng Shen, and Peter F. Thall. A modified adaptive lasso for identifying interactions in the cox model with the heredity constraint. *Statistics &amp Probability Letters*, 93:126–133, Oct 2014.

Survival Analysis by Noninvasive Imaging: A Quantitative Radiomic Approach

ORIGINALITY REPORT

4%
SIMILARITY INDEX

PRIMARY SOURCES

1	library.isical.ac.in:8080 Internet	43 words — 1%
2	export.arxiv.org Internet	41 words — 1%
3	www.duo.uio.no Internet	23 words — 1%
4	virtual.aapm.org Internet	17 words — < 1%
5	atrium.lib.uoguelph.ca Internet	14 words — < 1%
6	www.asems.mod.uk Internet	14 words — < 1%

EXCLUDE QUOTES ON
EXCLUDE BIBLIOGRAPHY ON
EXCLUDE SOURCES < 14 WORDS
EXCLUDE MATCHES < 14 WORDS