Non-linear Survival Analysis For Intrahepatic Cholangiocarcinoma Using Deep Learning with Radiographic Images

Katy Scott
School of Computing
Queen's University
Kingston, ON
15kls3@queensu.ca

Abstract

Intrahepatic cholangiocarcinoma (iCCA) is a cancer of the bile ducts within the liver with poor prognosis and minimal treatment options that is difficult to diagnose. Prognostic information determined from survival analysis could provide important insight for treatment development and general understanding of this disease. Deep learning methods applied to survival analysis have been shown to outperform both classical statistical methods (e.g., Cox proportional hazards models) and other machine learning techniques. These deep survival analysis models have been applied to a range of cancer types and input modalities. This work describes the development of a deep learning model that utilizes preoperative computed tomography (CT) images to predict recurrence-free survival in iCCA patients. Data has been retrospectively acquired from multiple institutions for patients who underwent a resection procedure. The final model achieved a Harrell's concordance index score of 0.54.

1 Introduction

Cholangiocarcinoma (CCA) is a cancer of the bile ducts that connect the liver, gallbladder, and small intestine. Subclassifications have been made based on the origin of the cancer; intrahepatic cholangiocarcinoma (iCCA) refers to those that start in the smaller branches within the liver, while perihilar and distal originate at the entrance to the liver and further down the duct closer to the small intestine, respectively.

Although rare, an increase in both incidence and mortality of CCAs has been observed worldwide in the past few decades, with 5-year survival rates being reported as ranging from 7-20% [3]. These low rates are a consequence of CCA being difficult to diagnose and treat effectively.

CCAs are typically asymptomatic in early stages [4]; 20-25% of diagnosed cases of iCCA arise from incidental findings [2]. iCCAs are also known to be architecturally heterogeneous in nature, likely due to the variation in cell function and morphology along the bile duct [27], which causes atypical clinical presentation and imaging [7]. Combined with a lack of screening strategies, diagnosis is often performed at advanced stages of the disease [7, 11].

As for treatment, tumour resection is currently the only potentially curative option [3, 6]. Moreover, percentage recurrence remains high after resection and only a small percentage of patients are eligible for the procedure [3].

Prognostic information could provide insight into patterns of disease progression and therefore treatment options. To date, no accurate non-invasive biomarkers have been identified for iCCA

CISC867 Deep Learning (Winter, 2021), Kingston, Canada.

diagnosis and prognosis estimations [3]. Accurate disease prognosis in general remains a prevalent challenge for physicians. Prediction factors considered for cancer patients can range from histological, clinical, and population-based data to environmental factors, family history, and age, among many others [19]. With the usage of high-throughput technologies, the magnitude of available data far exceeds that which could be accurately processed by a physician.

Various survival analysis techniques have been developed to improve on physician prediction, including Kaplan-Meier survival curves [15], the Cox proportional hazards model [8], and logistic regression [23]. Large amounts of available data make these statistical methods more difficult to apply without loss of information. Machine learning methods, however, have been presented as a strong contender to assist with this issue. In a 2015 review, Kourou et al. state that machine learning methods show a 15-25% improvement in accuracy of survival prediction over traditional methods [19]. It has also been suggested that machine learning might offer insight into prognostic and diagnostic investigation of CCAs [3].

Primitive forms of machine learning such as the artificial neural network showed success in predicting 5-year relapse of breast cancer over 20 years ago [9]. More recently, random survival forests and Bayesian frameworks have been applied to various survival prediction cases, including for lung cancer and coronary heart disease [5, 25]. The latest works, however, have begun to experiment with the use of deep learning (DL) methods, showing that they outperform not only classical statistical methods but other machine learning techniques as well.

This work investigates applying a deep convolutional neural network to the problem of recurrence-free survival (RFS) prediction for iCCA patients who underwent surgical resection. The model is evaluated with Harrell's concordance index [12].

This project is put forth as a proof of concept for the use of a non-linear survival analysis model as a means of prognosis prediction for iCCA. The results produced by the limited model that is presented are promising and motivate continued work to develop a fully-expressed tool that could aid both physicians in their diagnosis and treatment of the disease and in further research of iCCA.

The remainder of this paper covers the following: Section 2 provides a background on the techniques in use; Section 3 describes the process of data acquisition, preprocessing, the model architecture in use, and experimental methods and results; related works in survival analysis and analysis of iCCA are briefly reviewed in Section 4.

2 Background

This section reviews some fundamental concepts for modeling of linear and non-linear survival analysis.

2.1 Survival Analysis

Survival analysis is the task of estimating the time to a particular outcome. In relation to medical prognosis, this estimation may be the length of a hospital stay, overall survival time, or time to disease recurrence to name a few [23]. For this work, we are interested in the time between a patient's surgical resection and recurrence of iCCA or death.

There are five main components involved in survival analysis: the event of interest, the observation period, covariates, censoring, and a survival function. The event of interest is described in the previous paragraph. The observation period is the time over which data is collected from subjects being studied. In our case, the observation period begins on the day of a patients surgical procedure.

Covariates are the data collected for use in the survival analysis, which corresponds to the CT images for this work. These are what the survival prediction is based upon.

Censoring is a particularly important element, as it handles the cases when a patient can no longer be observed but an event has not yet occurred. Censored patients may have withdrawn from a study before the conclusion of the observation period or survived the length of the observation period. Censoring allows us to still use this data, even though we do not know if or when an event will occur in these samples.

The last factor is the survival function, which is the probability that the time, *T*, a subject will survive without an event occurrence will be longer than a time, *t*:

$$S(t) = Pr(T > t). \tag{1}$$

Another reformulation of the survival function is the hazard function, which is the probability that a subject will not survive an extra amount of time, δ given that have already survived up to a certain time, t. It is formulated as:

$$h(t) = \lim_{\delta \to 0} \frac{Pr(t \leq T < t + \delta \mid T \geq t)}{\delta}. \tag{2}$$

The greater the hazard, the greater the risk of disease recurrence.

Survival analysis models produce estimates for the survival or hazard function. An example of this is the Cox proportional hazards model.

2.2 Cox Proportional Hazards

The Cox proportional hazards (CPH) model is a multivariate semiparametric regression model that is commonly used to model the survival of a subject using their covariates [8]. The CPH model makes a couple of assumptions, the first being that the covariates must be linearly proportionally related such that the relative hazard over time must be constant. The second assumption is that the hazard function is composed of two non-negative functions: a baseline hazard function, $h_0(t)$, and a exponential risk function, $r = e^{\mathbf{x}_i^\top \beta}$. This form of the hazard function can be written as:

$$h_i(t) = h_0(t) e^{\mathbf{x}_i^\top \beta}, \tag{3}$$

where $\mathbf{x}$ is a feature vector containing the covariates and β is the vector of hazard model coefficients associated with each feature. An important note to make is that the baseline hazard function only depends on the time and is independent of the covariates, while the opposite is true for the risk function.

CPH estimates the risk function by tuning the β's to optimize the Cox partial likelihood function:

$$PL(\beta) = \sum_{i=1}^{n} \delta_i \left[\mathbf{x}_i^\top \beta - log\Big(\sum_{j \epsilon \mathbb{R}} e^{\mathbf{x}_i^\top \beta} \Big) \right] \tag{4}$$

This estimation is based on a linear log-risk function, hence the linear component of the proportional hazards condition. In the case of the data in use in this work, we cannot assume that this condition is met and therefore need an alternative way to perform survival analysis.

2.3 Non-linear Survival Analysis

A method posed by Faraggi and Simon in 1995 replaces the linear risk function estimation with the output of a simple feed-forward neural network [10]. Covariates are fed into the network, which then computes a linear combination of the nonlinear features it extracts to estimate the log-risk function. A visual representation of this can be seen in Figure 1. The only assumption made in this model is that the log-risk function is continuous.

Although this simple model was unable to outperform the linear CPH model, the same loss function can be used for more complex deep learning models. A selection of examples of this extension are reviewed in Section 4.

3 Method

This section describes the methodology for image data acquisition and preprocessing, model architecture, training and evaluation. Experimental results are presented and discussed, including suggestions for further improvement of the model in future work.

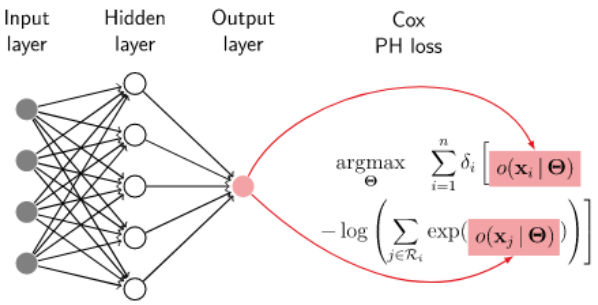

Figure 1: Example of how a feed-forward neural network with parameters θ can be used for survival analysis. The output of the neural network is plugged into the partial likelihood function (Eq. 4).

3.1 Data Acquisition

Images were acquired from patients with histologically confirmed primary iCCA that underwent curative intent resection at Memorial Sloan Kettering Cancer Center (USA) from 1993-2019 or Erasmus Medical Center (Netherlands) from 2005-2015. These studies were performed on a variety of scanner models all with multidetector CT (16–64 slices). All scans had maximal axial slice thickness of 5 mm. CT images were manually segmented to extract the tumour from the image. This generated tumour volumes in the form of MetaImage MetaHeader (MHD) and RAW files.

Recurrence-free survival after resection was calculated from the surgery date until documented recurrence or death in months. Patients alive and recurrence-free at last follow-up were censored and are identified by an RFS code of 0.

Twelve labelled samples were used from Erasmus and 173 from Memorial Sloan Kettering, for a total of 185 labelled tumour volumes. This dataset is not publicly available as it was shared with special permissions from these institutions for this work.

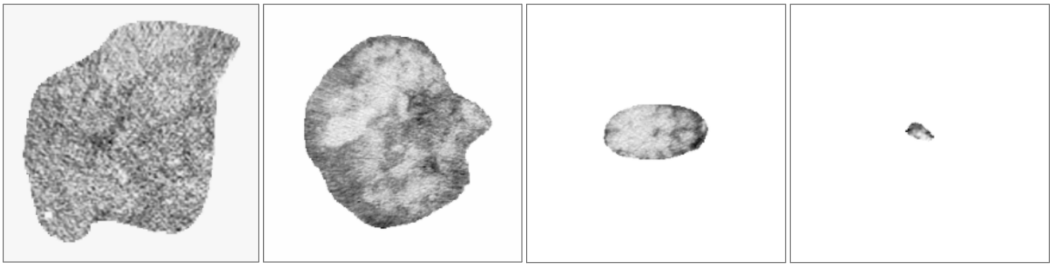

Figure 2: Four examples of segmented iCCA tumour image slices at the end of the preprocessing steps described in Section 3.2.

3.2 Image Preprocessing

Image preprocessing implementation was completed partially in MATLAB [21] and partially in Python [26].

For the first step, tumour volumes were processed only with those from the same institution. MHD files from a single institution were iterated over to find the maximum width and height in the set. Each image from the volume is then cropped based on these maximums and resized to a set dimension. This dimension can be set as the desired input size for the network and is set to 256 x 256 for this work. The background of the image is replaced with 0 values and the image is stored as a compressed binary file. After this process, 2888 image slices are available. Figure 2 shows some examples of the images at the end of preprocessing. This is what is fed to the neural network.

Labels are provided for each tumor volume, so a new CSV file is produced linking each binary slice file to the sample identifier and corresponding label. These labels include a binary RFS Code, indicating whether or not the sample has been censored, and a RFS Time value in months.

Image data and labels were then loaded into a Jupyter notebook for use in the network. Slices were split into training and evaluation sets based on their sample identifier and censoring status. This was purposefully done so that slices from the same patient were not split across training and evaluation, as this would nullify the effect of the evaluation set since the network would have seen data from that patient in training. The inclusion of censoring status in the consideration is to ensure the balance of censored and non-censored samples is similar in the training and evaluation sets.

3.3 Model Architecture and Implementation

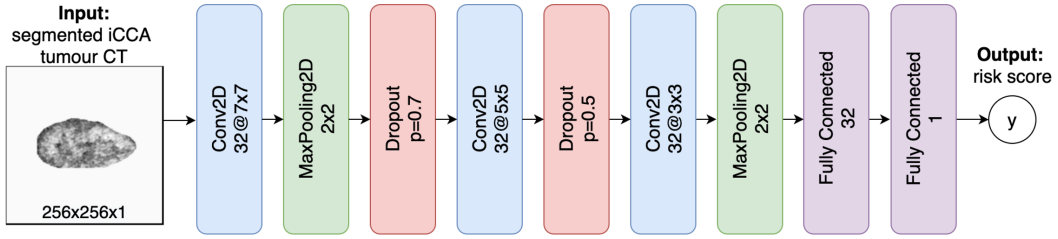

Figure 3: Overview of the KT6 model architecture. Convolution layers include the number of filters and kernel dimensions, max pooling layers include the kernel dimension, dropout layers include the dropout rate, and fully connected layers include the number of nodes.

Figure 3 shows a high level overview of the model architecture. The model consists of three convolutional blocks: the first contains a 2D convolution layer, with max pooling and a dropout layer with $p = 0.7$; the second, a 2D convolution and dropout layer with $p = 0.5$; and the last, a 2D convolution followed by a max pooling layer. The output from the last block is flattened and fed into a fully connected layer and is finally passed through a single linear node to output a predicted risk score.

All hidden layers use the "Scaled Exponential Linear Units" (SELU) activation function [18]. This variant of the rectified activation function is used in place of batch normalization to ensure zero mean and unit variance in each hidden layer to avoid exploding or diminishing gradients during training.

Dropout [14] layers were introduced as a regularization technique to help with the large imbalance in the dataset, as a majority of the samples are not censored.

The Adam optimizer [17] is used with default settings and a learning rate of 0.0003. The loss function to be minimized is the Cox partial likelihood function (Equation 4) with the linear components replaced by the outputs of the network parametrized by theta, $o(x_i|\theta)$, as described in Section 2.3:

$$PL(\beta) = \sum_{i=1}^{n} \delta_i \left[o(x_i|\theta) - log\Big(\sum_{j \epsilon \mathbb{R}} e^{o(x_i|\theta)} \Big) \right]. \tag{5}$$

The implementation for calculating this loss as part of the network is based upon a tutorial by Sebastian Pölsterl [24].

The model was trained over 100 epochs with a batch size of 64.

3.4 Model Evaluation

Harrell's concordance index[12] is used to measure model performance, as it is known to be the most accurate and suitable method for estimating prediction error and is a commonly used metric for survival prediction. The C-index measures how well a model predicts the ordering of event occurrence for subjects, which in this case is the recurrence of disease post-surgery. For context, a $c = 0.5$ is the average C-index of a random model, whereas $c = 1$ is a perfect ranking of death times [12, 16].

For comparison, I chose to use train another model using the DeepConvSurv model architecture described by Zhu et al. [30], but with the all the other hyperparameters kept the same as KT6. Training results for both my implementation of DeepConvSurv and KT6 are summarized in Table 1. Visualizations of these metrics during the training process can be seen in Figure 4.

Table 1: Summary of results for the KT6 and DeepConvSurv model training and validation.

Model	Training Loss	Validation Loss	C-index
DeepConvSurv	1.424	6.533	0.598
KT6	2.404	2.950	0.546

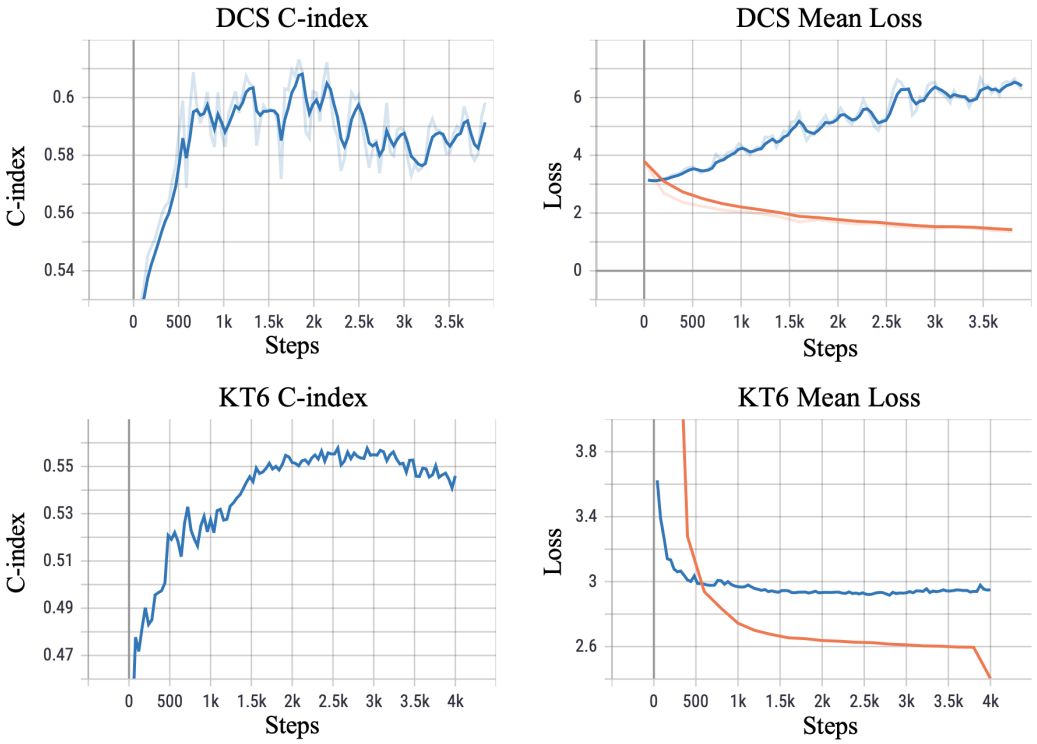

Figure 4: Training progress plots over 100 epochs for the DeepConvSurv (DCS) mode in the first row and the KT6 model in the second row. Values were saved every 200 steps during training. The first column shows the C-index score for the validation set. The second column shows the mean loss for the training set in orange and the validation set in blue.

Although my implementation of the DeepConvSurv model outperformed the KT6 for the C-index score, we can see from the mean loss graph in Figure 4 and reported validation loss value that the model appears to be overfitting. The KT6 model appears to be less affected this, potentially due to the introduction of the dropout layers. The validation loss in KT6 does not decrease much below 2.9 after 1500 steps, but the C-index continues to improve for another 1000 steps. The cause of the mismatch between the C-index and loss is not immediately clear and is left as a future exploration to be conducted in the continuation of this work. One possibility is that this loss metric is not suitable for comparison with the C-index metric and another should be introduced.

Another issue that might be causing the poor scores is the variance in the size of the tumour in the images. Looking at Figure 2, we can see what a difference there is between the first image example on the left and the significantly smaller tumour slice on the far right. Removing images with such small slices might be beneficial to training so the model is not learning from images that are more than 90% background. However, the image set is already small and the potential improvement in training gained by removing the small slice images could be cancelled out by the diminished number of images. Therefore, more data collection or image augmentation should be introduced to combat both

of these issues. This may also involve generating recurrence-free survival times for the augmented images, which is done by Li et al. by sampling from a Gaussian distribution with the original image RFS time, *T*, as the mean and $0.05 * T$ as the standard deviation [20].

Neither model performed as well as the original DeepConvSurv model [30] (C-index = 0.629), but for a first pass at implementation their results are promising, especially given the homogeneity of the iCCA tumours.

4 Related Works

This section reviews current work with the aim of answering two main queries. The first explores if deep learning methods can be successfully applied to survival analysis. The second asks if these DL methods will generalize specifically to iCCA and covers previous analysis done on the dataset in use in this work.

4.1 Deep Survival Analysis

DeepSurv, a Cox proportional hazards deep neural network model, was developed with the goals of providing personalized treatment recommendation and being a useful framework for medical research [16]. The model passes inputs through a fully-connected neural network with dropout that outputs a predicted log-risk function to be plugged back into a Cox model hazard function. DeepSurv, a random survival forest model, and a CPH model were all tested on both simulated and real survival data, as well as simulated and real treatment data. The real survival data included gene expression data and clinical features for breast cancer patients from the Molecular Taxonomy of Breast Cancer International Consortium. DeepSurv performed better than both the forest and CPH model on this dataset, achieving a C-index of 0.654. This is considered a foundational work for deep survival analysis that introduced a versatile model that could model risk effects of gene expression data.

Katzman et al. had suggested that the DeepSurv model could be extended with convolutional neural networks to predict risk with medical imaging. This extension was accomplished by Zhu et al. in their work on deep survival analysis on lung cancer data [30]. The model, aptly named DeepConvSurv, replaces the exponential part of a CPH model with a deep convolutional network to make survival predictions directly from pathological images. The model was tested on clinical and image data from 450 lung cancer patients and compared with a CPH model with Lasso feature selection, a CPH model with SuperPC feature transformation, a random survival forest model, and DeepSurv. When evaluated with 5-fold cross validation, DeepConvSurv outperformed the competition and achieved a C-index of 0.629.

In their work searching for prognostic biomarkers for high grade serous ovarian cancer (HGSOC), Wang et al. demonstrate the capability of unsupervised DL in survival prediction with CT images [28]. Their model extracts characteristics of HGSOC from preoperative CT images - one of the modalities suggested for use in this proposal. The network is comprised of a convolutional auto-encoder that takes in a 64 x 64 tumour image and outputs a 16-dimensional feature. A multivariate CPH regression is then used to build association between this DL feature and recurrence of HGSOC to predict a hazard score and RFS. When measured against the actual recurrence time, the CPH produced a C-index of 0.717 for training and validation scores of 0.713 and 0.694.

Further confirmation of image data being suitable for deep survival analysis comes from Li et al., who used a DL method to learn discriminative radiomic features from multi-modal imaging data for survival prediction of rectal cancer [20]. Their model is a more standard convolutional neural network, containing a series of convolution layers, a pooling layer, and concluding with fully connected layers leading to the final output layer. As with DeepSurv [16], the output is the estimated risk function from the proportional hazards function. The model was evaluated on a simulated dataset and pre-treatment PET and CT scans, both individually and in combination. In comparison with a CPH and random survival forest model, the DL model again came out on top with a C-index of 0.62 on the individual CT and 0.64 on the combined images.

4.2 Analysis of iCCA

Muhammad et al. utilized a deep clustering convolutional autoencoder to perform unsupervised subtyping of iCCA [22]. Prognostic prediction can be aided by cancer subtyping, which is challenging for rarer cancers, like iCCA, due to small patient cohorts.The model presented in this work clusters histopathological patterns of potential prognostic significance based on visual similarity. These clusters are used to train a CPH model, which showed a strong significance in RFS between patients in the same class. This work was performed on the same patient cohort as is presented in this proposal. Therefore, it is possible that similar clusters could be found and used for survival prediction in my model.

A subset of this dataset is used to present how preoperative CT images are associated with genetic pathways, disease-free and overall survival [1]. Two of the three imaging features shown to be significantly associated with increased risk of death were also associated with increased risk of recurrence. These were discovered using the CPH regression. Knowing that the deep survival methods presented above have outperformed CPH models suggests that these preliminary findings might be improved upon by the proposed project model. Connection between the image and genetic data was explored with Fisher's Exact test and the Wilcoxon Rank Sum, but no significant results were discovered. However, a study of another iCCA cohort did manage to find an association between distinct histomorphological features and mutation in *IDH1* using Fisher's Exact test [29].

These three works have shown that there exist associations in iCCA between histomorphological-based subtypes and RFS [22], CT images and RFS [1], and histomorphological image features and genomic features [29]. Along with the ideas developed in Section 4.1, a strong foundation exists to support the success of this work.

5 Conclusion

Intrahepatic cholangiocarcinoma is a difficult disease to diagnose and treat. The method developed in this work aimed to predict recurrence-free survival times for iCCA patients who had undergone surgical resection from segmented tumour CT images using a deep convolutional neural network for non-linear survival analysis. Results showed that the model requires further tuning, with a C-index score of 0.54 falling short of the scores achieved in comparable models seen in Section 4.

From the results of these other deep survival models we can see that there is plenty of room for improvement in the model presented in this work. Future work will include growing the data set, either by retrieving more image sets from MSK or Erasmus, performing image augmentation, or some combination of the two.

Another planned exploration is to experiment with incorporating pre-trained image classification models, such as ResNet101 [13], into the survival analysis model.

Lastly, non-linear survival analysis models are typically compared to Cox proportional hazards and random survival forest models. These require feature selection to be performed on the data, as the images contain too many pixels to be used as is in either analysis. This comparison will determine if the deep learning method being applied is having any impact on the prognostic predictions that are possible with this data type.

This work showcases a first step on the road to gaining new insight into intrahepatic cholangiocarcinoma prognosis with deep survival analysis.

Acknowledgments and Disclosure of Funding

This research is supported by the Vector Scholarship in Artificial Intelligence, provided through the Vector Institute. I would also like to thank my supervisors, Dr. Amber Simpson and Dr. Randy E. Ellis, for their guidance and support, as well as Dr. Travis Williams for his assistance in this project.

References

[1] Emily A. Aherne et al. "Intrahepatic cholangiocarcinoma: can imaging phenotypes predict survival and tumor genetics?" In: *Abdominal Radiology* 43.10 (2018), pp. 2665–2672. ISSN:

23660058. DOI: 10.1007/s00261-018-1505-4. URL: https://doi.org/10.1007/s00261-018-1505-4.

[2] Domenico Alvaro et al. "Cholangiocarcinoma in Italy: A national survey on clinical characteristics, diagnostic modalities and treatment. Results from the " Cholangiocarcinoma" committee of the Italian Association for the Study of Liver disease". In: *Digestive and Liver Disease* 43.1 (Jan. 2011), pp. 60–65. ISSN: 15908658. DOI: 10.1016/j.dld.2010.05.002.

[3] Jesus M. Banales et al. *Cholangiocarcinoma 2020: the next horizon in mechanisms and management*. Sept. 2020. DOI: 10.1038/s41575-020-0310-z. URL: https://doi.org/10.1038/.

[4] Jesus M. Banales et al. *Expert consensus document: Cholangiocarcinoma: current knowledge and future perspectives consensus statement from the European Network for the Study of Cholangiocarcinoma (ENS-CCA)*. May 2016. DOI: 10.1038/nrgastro.2016.51. URL: www.nature.com/nrgastro.

[5] James A. Bartholomai and Hermann B. Frieboes. "Lung Cancer Survival Prediction via Machine Learning Regression, Classification, and Statistical Techniques". In: *2018 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2018*. Institute of Electrical and Electronics Engineers Inc., Feb. 2019, pp. 632–637. ISBN: 9781538675687. DOI: 10.1109/ISSPIT.2018.8642753.

[6] Giovanni Brandi et al. "Genetic heterogeneity in cholangiocarcinoma: A major challenge for targeted therapies". In: *Oncotarget* 6.17 (2015), pp. 14744–14753. ISSN: 19492553. DOI: 10.18632/oncotarget.4539. URL: /pmc/articles/PMC4558112/%20/pmc/articles/PMC4558112/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4558112/.

[7] Y. S. Chong et al. "Differentiating mass-forming intrahepatic cholangiocarcinoma from atypical hepatocellular carcinoma using gadoxetic acid-enhanced MRI". In: *Clinical Radiology* 67.8 (Aug. 2012), pp. 766–773. ISSN: 00099260. DOI: 10.1016/j.crad.2012.01.004. URL: https://pubmed.ncbi.nlm.nih.gov/22425613/.

[8] D. R. Cox. "Regression Models and Life-Tables". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 34.2 (Jan. 1972), pp. 187–202. ISSN: 00359246. DOI: 10.1111/j.2517-6161.1972.tb00899.x. URL: http://doi.wiley.com/10.1111/j.2517-6161.1972.tb00899.x.

[9] Michelino De Laurentiis et al. "A Prognostic Model That Makes Quantitative Estimates of Probability of Relapse for Breast Cancer Patients". In: *Clinical Cancer Research* 5.12 (Dec. 1999), 4133 LP –4139. URL: http://clincancerres.aacrjournals.org/content/5/12/4133.abstract.

[10] David Faraggi and Richard Simon. "A neural network model for survival data". In: *Statistics in medicine* 14.1 (1995), pp. 73–82.

[11] Alejandro Forner et al. "Clinical presentation, diagnosis and staging of cholangiocarcinoma". In: *Liver International* 39.S1 (May 2019), pp. 98–107. ISSN: 1478-3223. DOI: 10.1111/liv.14086. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/liv.14086.

[12] Frank E. Harrell et al. "Evaluating the Yield of Medical Tests". In: *JAMA: The Journal of the American Medical Association* 247.18 (May 1982), pp. 2543–2546. ISSN: 15383598. DOI: 10.1001/jama.1982.03320430047030. URL: https://jamanetwork.com/.

[13] Kaiming He et al. "Deep residual learning for image recognition". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

[14] Geoffrey E Hinton et al. "Improving neural networks by preventing co-adaptation of feature detectors". In: *arXiv preprint arXiv:1207.0580* (2012).

[15] E. L. Kaplan and Paul Meier. "Nonparametric Estimation from Incomplete Observations". In: *Journal of the American Statistical Association* 53.282 (1958), pp. 457–481. ISSN: 1537274X. DOI: 10.1080/01621459.1958.10501452. URL: https://www.tandfonline.com/action/journalInformation?journalCode=uasa20.

[16] Jared L. Katzman et al. "DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network". In: *BMC Medical Research Methodology* 18.1 (Feb. 2018), p. 24. ISSN: 14712288. DOI: 10.1186/s12874-018-0482-1. arXiv: 1606.00931. URL: https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-018-0482-1.

[17] Diederik P Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (2014).

[18] Günter Klambauer et al. "Self-normalizing neural networks". In: *arXiv preprint arXiv:1706.02515* (2017).

[19] Konstantina Kourou et al. *Machine learning applications in cancer prognosis and prediction*. 2015. DOI: 10.1016/j.csbj.2014.11.005. URL: https://www.sciencedirect.com/science/article/pii/S2001037014000464.

[20] Hongming Li et al. "Deep convolutional neural networks for imaging data based survival analysis of rectal cancer". In: *Proceedings - International Symposium on Biomedical Imaging*. IEEE, 2019, pp. 846–849. ISBN: 9781538636404. DOI: 10.1109/ISBI.2019.8759301. arXiv: 1901.01449.

[21] *MATLAB version 9.9.0.1467703 (R2020b)*. The Mathworks, Inc. Natick, Massachusetts, 2020.

[22] Hassan Muhammad et al. "Unsupervised subtyping of cholangiocarcinoma using a deep clustering convolutional autoencoder". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 11764 LNCS. Springer, Oct. 2019, pp. 604–612. ISBN: 9783030322380. DOI: 10.1007/978-3-030-32239-7_67. URL: https://link.springer.com/chapter/10.1007/978-3-030-32239-7%7B%5C_%7D67.

[23] Lucila Ohno-Machado. *Modeling medical prognosis: Survival analysis techniques*. Dec. 2001. DOI: 10.1006/jbin.2002.1038.

[24] Sebastian Pölsterl. *Survival Analysis for Deep Learning Tutorial for TensorFlow 2*. 2020. URL: https://k-d-w.org/blog/2020/05/survival-analysis-for-deep-learning-tutorial-for-tensorflow-2/.

[25] Rajesh Ranganath et al. "Deep Survival Analysis". In: *Proceedings of the 1st Machine Learning for Healthcare Conference*. Vol. 56. Proceedings of Machine Learning Research. Northeastern University, Boston, MA, USA: PMLR, Aug. 2016, pp. 101–114. URL: http://proceedings.mlr.press/v56/Ranganath16.html.

[26] G. van Rossum. *Python tutorial*. Tech. rep. CS-R9526. Amsterdam: Centrum voor Wiskunde en Informatica (CWI), May 1995.

[27] Carlie S. Sigel et al. "Intrahepatic Cholangiocarcinomas Have Histologically and Immunophenotypically Distinct Small and Large Duct Patterns". In: *American Journal of Surgical Pathology* 42.10 (Oct. 2018), pp. 1334–1345. ISSN: 15320979. DOI: 10.1097/PAS.0000000000001118. URL: /pmc/articles/PMC6657522/%20/pmc/articles/PMC6657522/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6657522/.

[28] Shuo Wang et al. "Deep learning provides a new computed tomography-based prognostic biomarker for recurrence prediction in high-grade serous ovarian cancer". In: *Radiotherapy and Oncology* 132 (2019), pp. 171–177. ISSN: 0167-8140. DOI: https://doi.org/10.1016/j.radonc.2018.10.019. URL: https://www.sciencedirect.com/science/article/pii/S0167814018335436.

[29] Tao Wang et al. "Distinct histomorphological features are associated with IDH1 mutation in intrahepatic cholangiocarcinoma". In: *Human Pathology* 91 (2019), pp. 19–25. ISSN: 15328392. DOI: 10.1016/j.humpath.2019.05.002. URL: https://doi.org/10.1016/j.humpath.2019.05.002.

[30] X. Zhu, J. Yao, and J. Huang. "Deep convolutional neural network for survival analysis with pathological images". In: *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2016, pp. 544–547. DOI: 10.1109/BIBM.2016.7822579.