

Лаборатория больших данных и машинного обучения  
Южно-Уральский государственный университет,  
Челябинск, Россия

Мультимодальный метод глубокого обучения  
для прогнозирования выживаемости при раке легких

Presenter:

Dr Sachin Kumar

# Введение

- Рак легких является ведущей причиной смерти от рака как у мужчин, так и у женщин в мире, на его долю приходится 19,4%-27% всех смертей от рака
- В широком смысле рак легких можно классифицировать на две группы: немелкоклеточный рак легких (НМРЛ), составляющий 85%, и мелкоклеточный рак легких (МКРЛ), составляющий оставшиеся 15%, который имеет плохой прогноз
- Пятилетняя выживаемость при раке легких ниже, чем при многих других ведущих локализациях рака, таких как колоректальный рак (64,5%), рак молочной железы (89,6%) и рак предстательной железы (98,2%). Более половины людей с раком легких умирают в течение одного года после постановки диагноза
- Точная оценка стадии заболевания и времени выживаемости при раке легких имеет важное значение для принятия врачами решения об оптимальном плане и сроках лечения
- Lung cancer is the leading cancer killer in both men and women in the world, representing 19.4%–27% of all deaths from cancer
- It can be broadly classified into non-small cell lung cancer (NSCLC), counting for 85% and small cell lung cancer (SCLC), counting for the remaining 15%, which has a poor prognosis
- The lung cancer five-year survival rate is lower than many other leading cancer sites, such as colorectal (64.5%), breast (89.6%), and prostate (98.2%). More than half of people with lung cancer die within one year of being diagnosed
- Accurate assessment of the disease stage and survival time of lung cancer is essential in deciding the optimal plan and timing treatment for the clinicians

# Введение (продолжение...)

- В настоящее время значительная часть интерпретации медицинской информации выполняется медицинскими экспертами
- С точки зрения интерпретации изображений экспертами-людьми, в радиологии возникает множество диагностических ошибок.
- Ежегодно около 20 миллионов радиологических отчетов содержат клинически значимые ошибки
- Это ограничение в основном связано с субъективностью, усталостью эксперта, сложностью изображения, а также с большими вариациями у разных переводчиков
- Более того, 2/3 населения земного шара не имеют надлежащего доступа к специалистам в области радиологии; это составило бы 4,7 миллиарда человек
- Nowadays, a large fraction of interpretation of medical information is performed by medical experts
- In terms of image interpretation by human experts, a lot of diagnostic errors appear in radiology.
- Approximately 20 million radiology reports contain clinically significant errors each year
- This limitation is mainly due to the subjectivity, the fatigue of the expert, the complexity of the image, as well as the extensive variations across different interpreters
- Moreover, 2/3 of the world population lacks adequate access to radiology specialists; this would translate to 4.7 billion people

# Введение (продолжение...)

- Искусственный интеллект (ИИ) – многообещающий инструмент, который доказал свою эффективность в диагностических целях
- С быстрым развитием компьютерного зрения, основанного на глубоком обучении (такого как CNN), его способность распознавать изображения или диагностировать картинки даже превосходит возможности человека
- Более конкретно, это позволило моделям компьютерного зрения помочь врачам обнаружить подозрительные узелки в легких или определить местоположение узелка, оценить злокачественность всего легкого и спрогнозировать риск развития рака легких
- В некоторых случаях модели достигли конкурентных с врачами характеристик, а точность даже превосходит врачей
- Artificial intelligence (AI) is a promising tool that has shown its efficacy for diagnostic purposes
- With the rapid development in deep learning-based computer vision (such as CNN), its ability to recognize images or diagnose pictures even exceeds human ability
- More specifically, it enabled computer vision models to assist the doctors to detect suspicious pulmonary nodules or identify the location of the nodule, evaluate whole lung/pulmonary malignancy, and predict the risk of lung cancer
- Even in some cases, the models have reached competitive performance to doctors, and the accuracy even exceeds the doctors

# Введение (продолжение...)

- Однако в большинстве случаев простого рентгеновского снимка, компьютерной томографии недостаточно для диагностики заболевания; и врач должен рассмотреть другие варианты: история болезни пациентов, клинические данные, отчеты об анализах крови и др.
- Все эти типы источников информации о пациентах различны, и, следовательно, они называются разными модальностями
- Объединение этих методов для точной диагностики заболевания с помощью ИИ является сложной задачей
- В этом исследовании мы попытались разработать мультимодальный комбинированный подход для диагностики рака легких, используя как клинические данные, так и изображения 3D-компьютерной томографии
- However, in most of the cases, a simple X-ray image, CT scan is not enough to diagnose the disease; and a doctor has to consider other variants such as medical history of patients, clinical data, blood test reports etc.
- All these types of information sources about the patients is different and hence, they are called different modalities
- Combining these modalities together to accurately diagnose the medical disease using AI is a challenging task
- In this study, we attempted to develop a multimodal fusion approach to diagnose lung cancer using both clinical data and 3D CT scan images

# Набор данных

- В этой работе мы рассмотрели 422 пациента с НМРЛ (немелкоклеточным раком легких) из TCIA (Архив изображений рака), чтобы оценить предлагаемый метод
- Для этих пациентов доступны предварительная компьютерная томография, ручное определение радиационным онкологом трехмерного объема общей опухоли и данные о клинических исходах
- Соответствующие клинические данные также доступны в той же коллекции. Пациенты, для которых не были известны время выживаемости и статус события, были исключены из этой работы.
- In this work, we considered 422 NSCLC (Non-small Cell Lung Cancer) patients from TCIA (The Cancer Imaging Archive) to assess the proposed method
- For these patients pretreatment CT scans, manual delineation by a radiation oncologist of the 3D volume of the gross tumor volume and clinical outcome data are available
- The corresponding clinical data are also available in the same collection. The patients who had neither survival time nor event status were excluded from this work.

*\*Data Source:* K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, et al., “The cancer imaging archive (tcia): maintaining and operating a public information repository,” *Journal of digital imaging*, vol. 26, no. 6, pp. 1045–1057, 2013.

# Основные обозначения и предварительные знания

## А. Обозначения

Символ	Семантика
$r_i$	$i$ -е 3D радиологическое изображение
$c_i$	$i$ -я клиническая информация
$y_i$	Фактическое значение
$\hat{y}_i$	Предсказанное значение $y_i$
$e_i$	$i$ -е событие: 1 для пациентов без цензуры, 0 для пациентов с цензурой (неполная информация)
$I(x)$	1, если $x = True$ , иначе 0

## A. Notations

Symbol	Definition
$r_i$	$i^{\text{th}}$ 3D radiology image
$c_i$	$i^{\text{th}}$ clinical information
$y_i$	Actual value
$\hat{y}_i$	Predicted $y_i$
$e_i$	$i^{\text{th}}$ event, 1 for uncensored, 0 for censored patients (Incomplete Information)
$I(x)$	1 if $x = True$ else 0

# Основные обозначения и предварительные знания

## В. Мультимодальное глубокое обучение

- Мультимодальное глубокое обучение – это новая платформа, использующая глубокое обучение для изучения различных форматов, таких как текст, изображения и аудио
- В медицинских приложениях к каждому пациенту относятся различные типы данных, включая клиническую информацию, рентгенологические снимки, рекомендации врача, лекарства и многие другие
- Таким образом, когда данные поступают из разных источников, подход мультимодального глубокого обучения может помочь понять и извлечь больше полезной информации, а также обеспечить более точную диагностику

## В. Multimodal Deep Learning

- Multimodal deep learning is the novel framework using deep learning to learn from multiple modalities, such as text, images, and audio
- In medical applications, multiple types of data are related to each patient, including clinical information, radiology images, physician note, medication, to name a few
- Thus, when data comes from different sources, the approach of multimodal deep learning can help to understand and extract more useful information, and may provide more accurate diagnosis



# Основные обозначения и предварительные знания

## C. 3-D ResNet

- Остаточная нейронная сеть (ResNet) может справиться с проблемами исчезновения и взрыва градиента при более глубоком обучении нейронных сетей, особенно в области компьютерного зрения
- Основная концепция ResNet заключается в построении базового сетевого блока, в котором выходные данные суммируются с входными
- Чтобы обрабатывать ввод 3D-изображений, мы можем просто увеличить форму ядра с 2 измерений до 3 измерений (высота, длина и глубина) в слоях свертки
- The residual neural network (ResNet) can handle gradient vanishing or exploding problems in deeper neural networks training, especially in computer vision
- The core concept of ResNet is to construct a basic network block in which the output adds up with input.
- To handle 3D image input, we can simply increase kernel shape from 2 dimensions to 3 dimensions (height, length, and depth) in convolution layers

# Основные обозначения и предварительные знания

## D. Анализ Выживаемости

- Анализ выживаемости – это широко используемый метод анализа данных от времени до события (например, отсев учащихся, выживаемость при раке, госпитализация, рецидив заболевания и т.д.).
- Традиционные статистические методы анализа выживаемости обычно содержат три варианта моделирования функции выживаемости: непараметрические методы без предопределенного распределения кривой выживаемости, полупараметрические методы, такие как модель пропорциональных рисков Кокса (регрессия Кокса), которая применяется наиболее часто, и параметрические методы с предопределенным распределением (например, линейная регрессия, распределение Вейбулла).

## D. Survival Analysis

- Survival analysis is a widely used technique to analyze time to-event data (e.g., student dropout, cancer survival, admission to hospital, disease recurrence, etc.)
- Traditional statistical methods for survival analysis normally contains three options for modeling the survival function: non-parametric methods with no distribution of survival curve predefined semi-parametric methods such as the Cox proportional hazards model (Cox regression) which is most commonly applied, and parametric methods with distribution predefined (e.g., Linear regression, Weibull distribution).

# Основные обозначения и предварительные знания

## D. Анализ Выживаемости

- Из-за наличия подвергнутых цензуре данных о выживаемости (обычно подвергнутых правильной цензуре) стандартные оценочные индексы для регрессии, такие как среднеквадратичная ошибка (MSE) и  $R^2$ , не подходят для количественной оценки эффективности анализа выживаемости.
- Наиболее важным оценочным индексом является C-index\*, который может оценивать экземпляры без цензуры и подвергнутые цензуре экземпляры вместе

\* Индекс соответствия (С-индекс) – это широко используемый показатель в анализе выживаемости для оценки достоверности модели прогнозирования

$$C_{index} = \frac{\sum_{i,j} I(\hat{y}_i < \hat{y}_j \mid e_i = 1, y_i < y_j)}{\sum_{i,j} I(y_i - y_j \mid e_i = 1)}$$

## D. Survival Analysis

- Due to the existence of censored survival data (usually right censored), the standard evaluation indexes for regression, such as mean square error (MSE) and  $R^2$ , do not fit for quantifying the performance of survival analysis. The most important evaluation index is C-index\*, which can evaluate uncensored instances and censored instances together

\*The Concordance Index (C-index) is a commonly used metric in Survival Analysis to evaluate the goodness of a prediction model

C-index refers to the proportion of pairs whose predicted results are consistent with actual results among all patient pairs. For instance, the C-index is increased by 1% implies that if a survival analysis is made for a population of one million people, the correct prediction will increase the number of people by four thousand

# Основные обозначения и предварительные знания

## D. Анализ Выживаемости

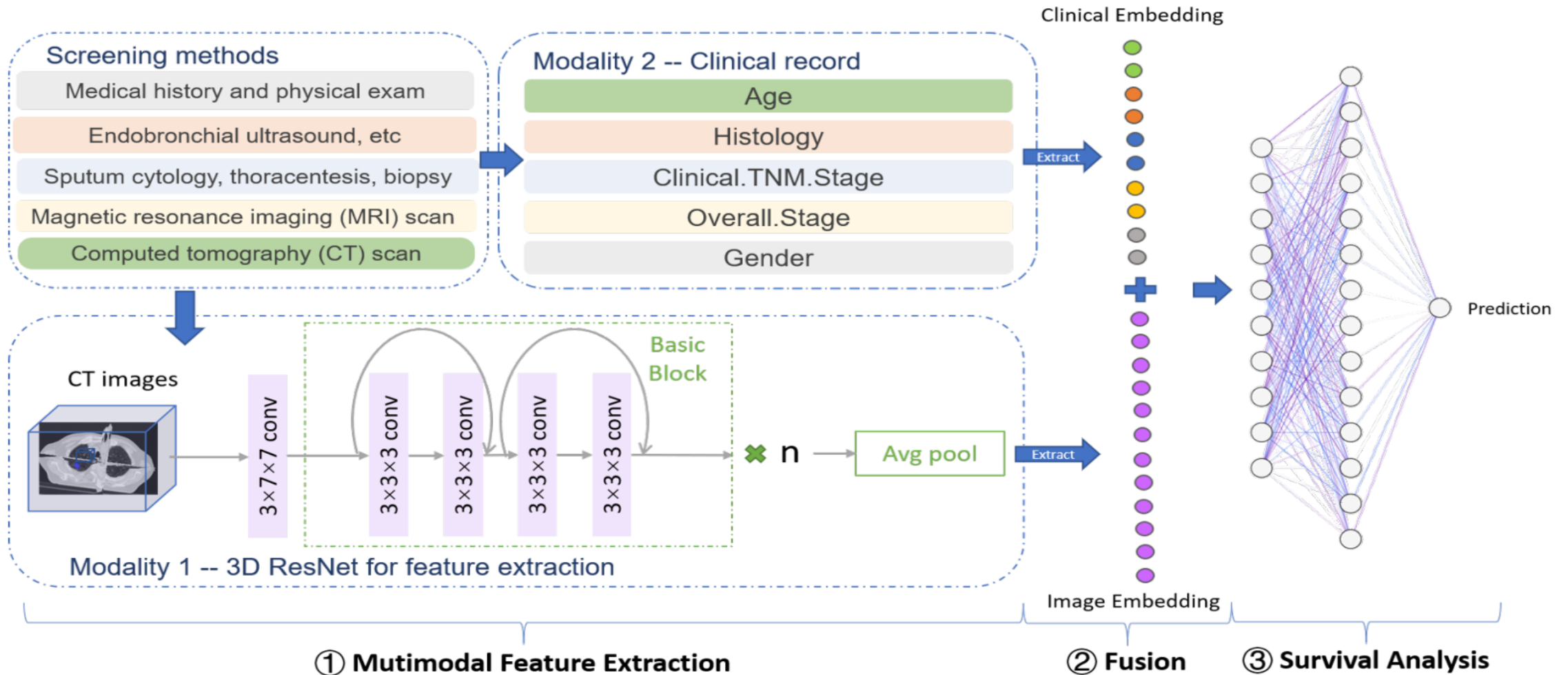
- Кроме того, в качестве задачи регрессии мы также используем среднюю абсолютную ошибку (MAE) по экземплярам без цензуры для оценки наших экспериментов

## D. Survival Analysis

- Besides, as a regression problem, we also use the mean absolute error (MAE) over uncensored instances to evaluate our experiments

$$MAE = \frac{1}{\sum_i I(e_i = 1)} \sum_i^N (e_i |y_i - \hat{y}_i|)$$

# Предложенный фреймворк



- (1) Во-первых, он использует 3D-ResNet в сочетании с обычными сетями для мультимодального извлечения признаков;
- (2) Затем он использует метод простого слияния признаков (early fusion) для мультимодального слияния;
- (3) Наконец, на этапе принятия решения простая нейронная сеть предназначена для прогнозирования выживаемости.

# Мультимодальное извлечение признаков

## 1a) Извлечение объектов КТ-изображения с помощью 3D-ResNet:

- Мы предложили 3D-ресети в качестве нашей сетевой структуры для низкоуровневого изучения особенностей изображений
- Чтобы устранить проблемы типа исчезающих и взрывающихся градиентов в очень глубокой нейронной сети, мы добавляем дополнительные короткие соединения в нашу модель. Поскольку наш набор данных относительно невелик по сравнению с другими общими наборами данных для распознавания изображений, мы используем методы аугментации данных перед импортом данных в модель

## 1b) Извлечение признаков из истории болезни:

- Скрытые состояния клинических данных вводят клинические данные в сеть отдельно, что направлено на сбор информации о выживаемости, имеющейся в клинических данных
- Для извлечения признаков из клинических данных была предложена нейронная сеть с двумя скрытыми слоями

## 1a) CT Image Feature Extraction with 3D-ResNet:

- We proposed 3D-ResNets as our network structure for the low-level image feature learning
- To eliminate the vanishing and exploding gradients types of problems in the very deep neural network, we add extra shortcut connections in our model
- Since our dataset is relatively small compared to other general datasets for image recognition, we use data augmentation techniques before importing the data into the model

## 1b) Clinical Record Feature Extraction:

- The clinical embedding layers introduce clinical data to the network separately, aims to capture the survival information indicated within clinical data
- For clinical data, a neural network with two hidden layers was proposed to extract features

# Мультимодальное извлечение признаков

- На этапе объединения признаков мультимодальные признаки с КТ-изображений и клинических записей трудно непосредственно объединить. Это связано с тем, что функции из разных модальностей имеют разные масштабы или статистические свойства.
- Таким образом, чтобы решить эту проблему, мы применили метод пакетной нормализации (BN) для корректировки среднего значения и дисперсии извлеченных признаков в каждом режиме перед процедурой слияния.
- Для заданных признаков  $z_1, z_2, \dots, z_m$  пакета,  $\hat{z}_i$  вычисляется следующим образом:
- During the feature fusion stage, multimodal features from CT images and clinical records are difficult to be directly fused together.
- This is due to the features from different modalities have different scales or statistical properties.
- Thus, to solve this problem, we applied the Batch Normalization (BN) technique to adjust the mean and variance of extracted features in each modality before the fusion procedure.
- Given the features  $z_1, z_2, \dots, z_m$  over a batch,  $\hat{z}_i$  is calculated as:

$$\hat{z}_i = \gamma_i \frac{z_i - \mu_i}{\sqrt{\sigma_i^2 + \epsilon}} + \beta_i$$

где  $\gamma_i, \beta_i$  – параметры обучения,  $\mu_i$  – среднее значение  $z_i$  в пакете,  $\sigma_i$  – стандартное отклонение  $z_i$  в пакете,  $\epsilon$  – параметр-бесконечно малое, например,  $10^{-8}$

where  $\gamma_i, \beta_i$  are the parameters to be learned,  $\mu_i$  is the mean value of  $z_i$  over the batch,  $\sigma_i$  is the standard deviation of  $z_i$  over the batch,  $\epsilon$  is set to a very small number such as  $10^{-8}$

# Анализ выживаемости

- В этом модуле для удовлетворения требований, предъявляемых к конкретной задаче, может быть использована любая модель анализа выживаемости
- В этой работе мы определяем время выживания в качестве метки и используем нейронную сеть с одним скрытым слоем и одномерным выходным слоем для настройки регрессии и общего удобства оптимизации
- Чтобы достичь одинакового масштаба между базовой истинностью и выходными данными в нашей модели, сигмовидная функция используется в качестве метода нормализации для обработки выходных данных
- In this module, to meet the requirements for a specific problem, any survival analysis model can be used
- In this work, we define the survival time as the label, and use the one hidden layer neural network with one dimension output layer for the regression setting and overall optimization convenience
- To achieve the same scale between ground truth and the output in our model, the Sigmoid function is used as a normalization technique to process the output

$$\hat{y}_i = \frac{1}{1 + e^{-f(x_i)}}$$



## Анализ выживаемости

- В качестве целевой функции используется функция потерь MSE и штрафной термин регуляризации L2, который определяется как:
- MSE loss function and L2 regularization penalty term as the objective function is used, which is defined as:

$$\text{minimize } L = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 + \lambda \sum \omega_j^2$$

где  $\omega_j$  – параметр модели,  $\lambda$  – штрафной коэффициент

where  $\omega_j$  is the model parameter,  $\lambda$  is the penalty coefficient

# Эксперименты и результаты

- Мы провели обширные эксперименты на пациентах с НМРЛ из набора TCIA, чтобы подтвердить эффективность предложенного метода с помощью нескольких самых современных методов с точки зрения точности прогнозирования продолжительности жизни каждого пациента
- Кроме того, мы также оценили результат прогнозирования по C-индексу. После этого мы проводим несколько экспериментов по удалению различных структур сети, чтобы определить наилучшую структуру

## Data Preprocessing (CT Images)

- Для КТ-изображений мы изменили размер необработанных данных, представляющих собой трехмерный объем первичной общей опухоли, на  $96*96*8$ . После этого мы преобразуем линейность диапазона в  $[0,1]$
- Затем, чтобы предотвратить переобучение, мы выполняем аугментацию данных, которая включает в себя три метода: поворот, замену местами и переворот, тогда мы получили  $422*8 = 3376$  образцов, среди которых было  $373*8 = 2984$  образца без цензуры и  $49*8 = 392$  образца с цензурой

- We conducted extensive experiments based on NSCLC patients from TCIA to validate the performance of the proposed method with several state-of-the-art methods in terms of the prediction accuracy for the survival time for each patient
- Besides, we also evaluated the prediction result by C-index. Afterward, we perform several ablation experiments regarding different network structures to determine the best structure

## Data Preprocessing (CT Images)

- For CT images, we resized the raw data which is the 3D volume of the primary gross tumor volume into  $96 * 96 * 8$ . After that, we transform the range linearity into  $[0,1]$
- Then, to prevent overfitting, we perform data augmentation which includes three methods: rotate, swap, and flip, then we got  $422 * 8 = 3376$  samples, among which there were  $373 * 8 = 2984$  uncensored samples and  $49 * 8 = 392$  censored samples.

# Эксперименты и результаты

## Предобработка (клинические данные)

- Клинические данные содержат как категориальные, так и некатегориальные данные. Во-первых, недостающее значение является распространенной проблемой в медицинских данных и может создавать трудности для анализа и моделирования данных
- В наборе данных категория "возраст" содержала несколько пропущенных значений. После наблюдения за данными было обнаружено, что возраст пациентов в наборе данных близок друг к другу. Таким образом, вычисляется среднее значение и заполняет недостающие значения
- После этого, используется однократный кодер для кодирования категориальных данных в числа, что позволяет сделать представление категориальных данных более выразительным
- Затем мы используем метод минимаксную и z-нормализацию для выполнения нормализации данных признаков, таких как возраст и время выживания

## Data Preprocessing (Clinical Data)

- Clinical data contains both categorical data and non-categorical data. Firstly, missing value is a common problem in medical data and may pose difficulties for data analyzing and modeling
- In the dataset, the "age" category contained a few missing values. After observing the data, It was found that the age of patients in the dataset is close to each other. Thus, the mean value is imputed and filled into the missing values
- Afterward, to fit into the model, one-hot encoder to encode categorical data into numbers is used, which allows the representation of categorical data to be more expressive
- Then, we use the min-max feature scaling method and standard z-score method to perform data normalization, such as age and survival time

## Предобработка (клинические данные)

Для входного данного  $x$  результат нормализации с помощью минимаксного масштабирования равен

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

результат z-нормализации вычисляется как:

$$x' = \frac{x - mean(x)}{std(x)}$$

где  $std$  - стандартное отклонение. Для одного пациента с множественными опухолями мы выбрали первичный общий объем опухоли ("GTV-1") для обработки в нашей работе, в то время как другие опухоли, такие как вторичные объемы опухоли, обозначенные как "GTV-2", "GTV-3", и это лишь некоторые из них, которые иногда присутствовали, не рассматривались

## Data Preprocessing (Clinical Data)

For input  $x$ , the min-max feature scaling method's output is:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

And the standard score method's output is:

$$x' = \frac{x - mean(x)}{std(x)}$$

where  $std$  is the standard deviation. For a single patient with multiple tumors, we selected the primary gross tumor volume ('GTV-1') to be processed in our work, while other tumors such as secondary tumor volumes denoted as 'GTV-2', 'GTV-3' to name just a few, which were occasionally present, were not considered

## Параметры экспериментов

- Мы обучили и оценили метод на наборе данных NSCLC-Radiomic после 5-кратной перекрестной валидации с разделением на уровни пациентов
- Мы разделили набор данных на обучающие, валидационные и тестовые данные в соотношении 6:2:2 соответственно
- Для настройки гиперпараметров, таких как штрафной коэффициент, мы используем набор данных проверки для точной настройки и получения оптимизированных гиперпараметров
- В процессе обучения мы использовали в общей сложности 200 эпох с Адамом в качестве оптимизатора.

## Experimental Setup

- We trained and evaluated the method on the NSCLC-Radiomic dataset following 5-fold cross-validation with the patient-level split
- We divided the dataset into training, validation, and testing data into 6:2:2 respectively
- For hyper parameters tuning such as the penalty coefficient, we use the validation dataset to fine-tune and get the optimized hyper parameters
- In the training process, we used 200 epochs in total with Adam as the optimizer.

## Параметры экспериментов

- Параметр размера пакета был установлен равным 64. Начальная скорость обучения устанавливается равной 0,001, а затем скорость обучения снижается на 0,5 каждые 40 эпох
- Поскольку в качестве метки мы используем время выживания, а не совокупный риск, в процессе обучения и валидации мы используем только данные без цензуры для точного расчета времени выживания и целевой функции
- Далее, в процессе тестирования мы используем все данные для оценки C-индекса и данные без цензуры для оценки MAE

## Experimental Setup

- The batch size parameter was set as 64. The initial learning rate is set as 0.001, and then the learning rate is decayed by 0.5 in every 40 epochs
- Since we use survival time as the label, not cumulative hazard; In the training and validation process, we only use the uncensored data for precise survival time and objective function calculation
- Further, in the testing process, we use all data for C-index evaluation and uncensored data for MAE evaluation

## Исследование Абляции

Чтобы найти оптимальную сеть для нашей задачи, мы рассмотрим проведение экспериментов по удалению, основанных на следующих четырех аспектах сетевой архитектуры:

- Как глубина ResNet влияет на производительность?
- Какая 3D-структура является лучшей?
- Превосходят ли мультимодальные методы по эффективности методы с одной модальностью?
- Каково наилучшее соотношение между данными изображения и клиническими данными на этапе слияния?
- Следует ли добавлять скрытые слои на этапе анализа выживаемости?

## Ablation Study

To find an optimal network for our problem, we consider performing **ablation experiments** based on the following four aspects of network architecture:

- How the depth of ResNet affect the performance? Which 3D structure is the best?
- Whether multiple modalities outperform single modality?
- What is the best ratio set between image data and clinical data in the fusion stage?
- Whether hidden layers should be added in the survival analysis stage?

## Исследование абляции

- Во-первых, проводится оценка различной глубины Resnet и того, являются ли несколько методов лучше, чем один
- Используя один скрытый слой в сети анализа выживаемости и установив соотношение признаков множественных модальностей на стадии слияния равным 512:27, были протестированы четыре различные структуры
- Таблица показывает, что наша модель структуры ResNet34 с несколькими модальностями достигла наилучшей производительности
- Все экспериментальные результаты подтверждают наше предположение и показывают эффективность использования нескольких методов анализа выживаемости

Структура ResNet	КТ-изображения		Мультимодальность	
	Ф-ция потерь	C-Index	Ф-ция потерь	C-Index
ResNet18	0.1023	0.5782	0.0847	0.6287
ResNet34	0.0975	0.5942	<b>0.0757</b>	<b>0.6490</b>
ResNet50	0.1023	0.5804	0.0760	0.6375
ResNet101	0.1071	0.5660	0.0795	0.6142

## Ablation Study

- Firstly, the evaluation of different depths of Resnet and whether multiple modalities are better than a single modality are conducted
- By using one hidden layer in the survival analysis network and fixing the ratio of multiples modalities features in the fusion stage to 512:27, four different structures were tested.
- Table shows that our model of **ResNet34** structure with multiple modalities achieved the best performance. All experimental results support our assumption and show the effectiveness of using multiple modalities in survival analysis

ResNet Structure	CT images		Multi-modality	
	Loss	C-Index	Loss	C-Index
ResNet18	0.1023	0.5782	0.0847	0.6287
ResNet34	0.0975	0.5942	<b>0.0757</b>	<b>0.6490</b>
ResNet50	0.1023	0.5804	0.0760	0.6375
ResNet101	0.1071	0.5660	0.0795	0.6142



## Исследование абляции

- Более того, используя наилучшую структуру ResNet34, мы делаем еще один шаг вперед, чтобы наблюдать эффекты изменения соотношения между модальностями со скрытым слоем и без него
- Это можно сделать, изменив количество перцептронов для объектов изображения и объектов, не связанных с изображением
- Список результатов сравнения в таблице показывает, что наилучшая производительность достигается при установке соотношения между изображениями и режимами без изображений равным 25:27
- Кроме того, добавление скрытых слоев в модуль анализа выживаемости может еще больше повысить производительность

Разбиение Image:Non Image	Скрытые слои		Без скрытых слоев	
	Ф-ция потерь	C-Index	Ф-ция потерь	C-Index
512:27	0.00757	0.6490	0.0760	0.6376
100:27	0.0745	0.6512	0.0755	0.6421
25:27	<b>0.0739</b>	<b>0.6580</b>	0.0761	0.6450
5:27	0.0765	0.6403	0.0793	0.6215

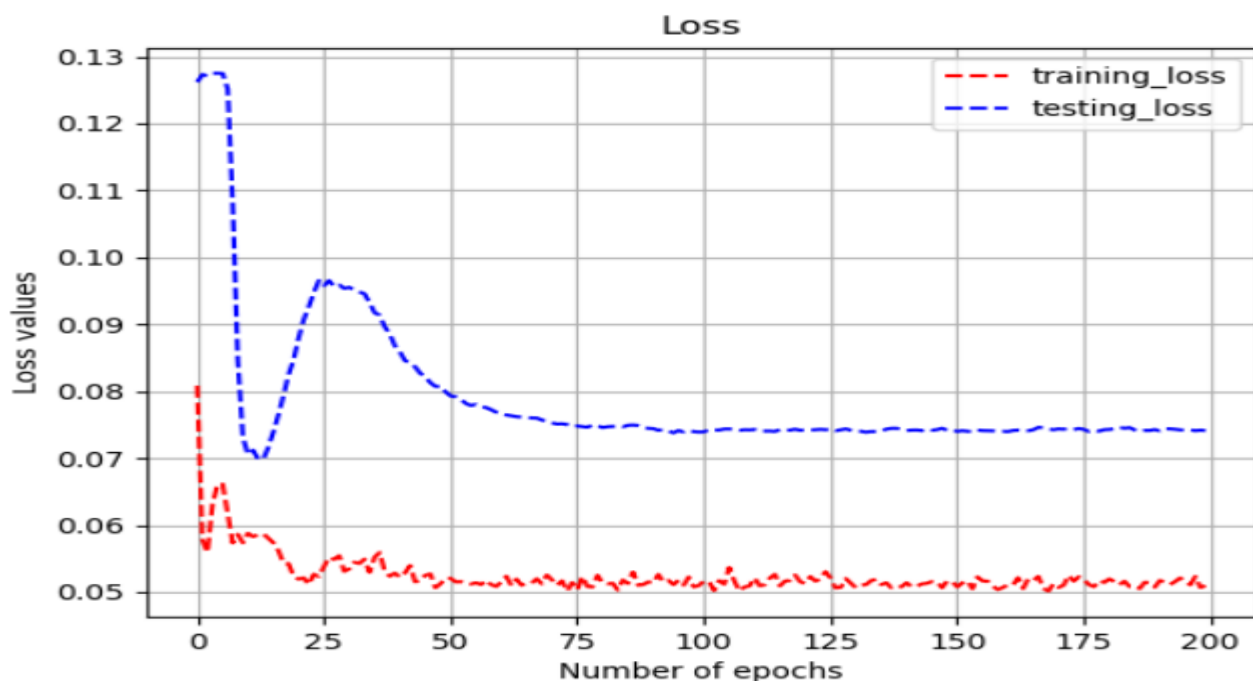
## Ablation Study

- Moreover, using the best structure **ResNet34**, we step further to observe the effects of changing the ratio between modalities with and without the hidden layer
- It can be done by changing the number of perceptrons for image and non-image features
- The comparison results list in Table shows that setting the ratio between images and non-image modalities as 25:27 achieves the best performance
- Besides, adding hidden layers in the survival analysis module can further improve the performance

Image: Non Image Split	Hidden Layer		No Hidden Layer	
	Loss	C-Index	Loss	C-Index
512:27	0.00757	0.6490	0.0760	0.6376
100:27	0.0745	0.6512	0.0755	0.6421
25:27	<b>0.0739</b>	<b>0.6580</b>	0.0761	0.6450
5:27	0.0765	0.6403	0.0793	0.6215

## Дальнейшее обсуждение

После экспериментов по удалению наилучшая архитектура с использованием структуры ResNet34, устанавливающая соотношение между несколькими модальностями как 25:27 с одним скрытым слоем, превосходит другие структуры в этой работе



## Further Discussion

After Ablation experiments, the best architecture of using **ResNet34** structure, setting the ratio between multiple modalities as **25:27** with one hidden layer outperform the other structures in this work

Результаты в ср. с линией отсчета

Модель	MAE	$C_{index}$
Cox-time	0.183	0.6152
Cox regression	0.204	0.6009
CoxCC	0.183	0.6120
PC-Hazard	0.191	0.6094
DeepHit	0.183	0.6133
<b>Предложенная модель</b>	<b>0.162</b>	<b>0.6580</b>

## Дальнейшие эксперименты

Несмотря на это, предложенная модель способна превзойти другие существующие статистические модели и модели машинного обучения.

Можно ожидать некоторых улучшений при внедрении следующих подходов:

- Промежуточное и позднее мультимодальное слияние может быть внедрено для оценки эффективности и усовершенствования существующего подхода
- Модуль анализа выживаемости может быть заменен существующей моделью анализа выживаемости, такой как Cox-time, глубокое обследование, чтобы улучшить эффективность предлагаемого подхода
- Можно собрать больше данных для оценки эффективности существующего подхода

## Further Planned Experiments

Although, the proposed model is able to outperformed other existing statistical/ML models. Some improvements can be expected by introducing the following approaches:

- The Intermediate and Late multimodal fusion can be introduced to evaluate the effectiveness and improvement in the existing approach
- The survival analysis module can be replaced with existing survival analysis model such as Cox-time, deepsurv to improve the performance of proposed approach
- More data can be collected to evaluate the performance of existing approach

Thank You!!  
Questions & Suggestions are Welcome!