

КЛАССИФИКАЦИЯ МУЛЬТИМОДАЛЬНЫХ ДАННЫХ О ЗАБОЛЕВАНИЯХ ЛЕГКИХ НА ОСНОВЕ ПОЗДНЕГО СЛИЯНИЯ МОДАЛЬНОСТЕЙ

© 2024 О.Н. Иванова, С. Кумар, М.Л. Цымблер, Е.В. Иванова

Южно-Уральский государственный университет

(454080 Челябинск, пр. им. В.И. Ленина, д. 76)

E-mail: onivanova@susu.ru, kumars@susu.ru, mzym@susu.ru, elena.ivanova@susu.ru

Поступила в редакцию: 21.09.2023

С развитием аппаратных технологий высококачественные рентгеновские снимки стали доступны для диагностики заболеваний легких с помощью специалистов-радиологов. Однако процесс диагностики занимает много времени и зависит от наличия в медицинском учреждении специалистов соответствующего профиля. В то же время информация о пациенте может включать не только рентгеновские снимки грудной клетки разного качества, а также результаты медицинских анализов, записи и предписания врача, сведения о приеме лекарств и другие. В данном исследовании предложена модель классификации легочных заболеваний на основе мультимодальных данных о клинических исследованиях пациентов и рентгенографических изображений. При подготовке данных использованы различные методы генерации искусственных образцов как для изображений, так и для табличных данных о результатах лабораторных исследований. Предложен метод установления соответствия для сгенерированных образцов между модальностями. Предложенная мультимодальная модель имеет архитектуру позднего слияния. Проведены эксперименты на наборах данных с одной и двумя модальностями. Предложенная модель показала точность на 5.5% выше, чем модели, основанные на одной модальности (91.3% против 86.11% на наборе данных из 1 156 пациентов).

Ключевые слова: мультимодальные данные, заболевания легких, глубокое обучение, позднее слияние.

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Иванова О.Н., Кумар С., Цымблер М.Л., Иванова Е.В. Классификация мультимодальных данных о заболеваниях легких на основе позднего слияния модальностей // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2024. Т. 13, № 1. С. 74–86. DOI: 10.14529/cmse240105.

Введение

В классификации ВОЗ число болезней, связанных с пневмонией и гриппом, составляет около 200 [1, 2]. Визуализация — один из возможных подходов к диагностике легочной патологии, в котором рентгенограммы используются для выявления инфекций легких [3], однако такая диагностика требует участия эксперта. Недостаточная квалификация врача или аномалии медицинских рентгенографических изображений, выполненных на различном оборудовании, приводит к потере времени при первичной диагностике. В настоящее время глубокие нейронные сети повсеместно внедряются в медицинскую практику. Во многих работах продемонстрирован положительный эффект применения глубоких нейронных сетей и машинного обучения в решении задач автоматизированного распознавания различных патологий, в том числе заболеваний легких [4, 5]. Одна из первых моделей для обнаружения пораженных легочных узлов была разработана в конце 1980-х годов, но оказалась недостаточной для выявления критических заболеваний легких на ранних стадиях. С развитием искусственного интеллекта исследователи провели ряд исследований, направленных на улучшение диагностики заболеваний легких. Однако медицинская экспертиза очень важна для понимания и диагностики критических заболеваний. Методы глубокого

обучения позволили разработать и обучить модель считать и понимать изображения легких и диагностировать, поражены они или нет. Однако качество любого прогноза модели глубокого обучения в значительной степени зависит от размера и качества данных. Далее приведены недавние исследования, иллюстрирующие состояние исследований в области классификации и прогнозирования заболеваний легких. В работе [6] предложена 3D глубокая сверточная сеть со стратегиями многомасштабного прогнозирования для обнаружения легочных узлов по сегментированным изображениям. Однако в данном исследовании не удалось классифицировать типы заболеваний, подход к прогнозированию применяется только для небольших легочных узелков. Сверточная нейронная сеть предложена в работе [7] для снижения частоты ложноположительных результатов при классификации набора данных о легочных узелках на наборе данных LUNA16. Этот метод позволяет анализировать характер изображений компьютерной томографии только для того, чтобы снизить вероятность неправильного диагноза. Более быстрый метод R-CNN был разработан в работах [7, 8] для обнаружения пораженных легочных узлов, а также снижения частоты ложноположительных результатов. Данная модель показала многообещающие результаты для обнаружения объектов. Слияние архитектуры глубокой сверточной сети и двухпутной сети (DPN, Dual Path Network) было предложено в работе [9] для классификации и извлечения признаков конкреций (минеральных образований). Расположение нескольких патчей с фильтром Франки было использовано в работе [10] для повышения производительности обнаружения легочного узла на рентгеновских снимках легких. Тем не менее эта модель показала чувствительность 94% при коэффициенте ложноположительных результатов 15.1. Применение алгоритмов искусственного интеллекта (ИИ) в классификации легочных заболеваний было предложено в работе [11] с учетом современного уровня классификации рентгенографии грудной клетки и ее анализа. В этой работе авторы создали набор данных, известный как ChestX-ray8, в котором 32 717 рентгеновских снимков принадлежат уникальным пациентам. В работе [12] авторы разработали глубокие сверточные сети для проверки результатов по этим данным о легких и добились многообещающих результатов. Набор данных ChestX-ray8 также адаптирован для использования для многоклассовой классификации заболеваний легких [12]. В работе [12] была предложена модель глубокого обучения для прогнозирования рака легких и пневмонии, предлагающая два метода глубокого обучения. Первоначально они использовали модифицированную модель AlexNet для диагностики рентгенографии грудной клетки. Более того, в модифицированной модели AlexNet был реализован метод опорных векторов (SVM, Support Vector Machine) при решении задачи классификации [13, 14].

В данной работе нами были решены следующие задачи, направленные на получение эффективного инструмента для борьбы с заболеваниями легких, планирования и принятия ответных мер для общественного здравоохранения в Челябинской области и других регионах России:

- провести эксперименты по достижению желаемой точности бинарной классификации наличия заболевания легких на разработанном наборе данных при обработке отдельных модальностей с помощью обучения различных предобученных моделей;
- разработать метод установления соответствия искусственно сгенерированных данных разных модальностей при позднем слиянии;
- разработать архитектуру и реализовать многомодальную модель классификации легочных заболеваний на основе позднего слияния;

- сравнить правильность предложенной многомодальной модели с результатами классификации на одиночных модальностях.

Статья организована следующим образом. В разделе 1 описан подход к слиянию мультимодальных данных. Раздел 2 содержит краткое описание набора данных, использованного в исследовании. В разделе 3 рассмотрены методы, с помощью которых проведено исследование. Результаты исследования и их обсуждение приведены в разделе 4. Заключение подводит итоги исследования.

1. Слияние мультимодальных данных

Слияние является одной из главных задач и вызовов, которые стоят перед исследователем мультимодальных данных. Принятие решений в процессе детектирования медицинских аномалий, классификации заболеваний и прогнозе лечения основывается на изучении различных источников лабораторных исследований — рентгенологических снимков, результатов исследования крови, томографических и многих других типов исследований. Многомодальные модели позволяют объединить данные разной природы различными способами: конкатенация, построение метамоделей, обучающейся после и во время добавления новых модальностей, мэшинг и др. Многомодальные данные — это данные, которые представлены в нескольких модальностях: текст, изображения, аудио и видео. Слияние многомодальных данных — это процесс объединения данных из разных модальностей для получения более полной и точной информации. Существует три основных подхода к слиянию многомодальных данных: раннее, позднее и промежуточное. При раннем слиянии входные данные из разнородных источников объединяются перед обработкой сетью глубокого обучения. Это самый простой подход, но он может быть неэффективен, если модальности неоднородны. При позднем слиянии сначала обучаются отдельные модели для каждой модальности. Затем результаты каждой модели объединяются для принятия решения. Этот подход более эффективен, чем раннее слияние, для обработки разнородных модальностей. При промежуточном слиянии данные из разных модальностей объединяются на промежуточных этапах обработки, принимая во внимание скрытые представления (embed) не только последних, но и внутренних слоев сетей различных модальностей. Этот подход требует значительных усилий по проработке архитектуры моделей и предназначен для обработки сильно разнородных модальностей.

2. Набор данных

Хорошо подготовленный, четко маркированный набор данных является ключевым фактором успеха высокоточного прогнозирования и классификации заболеваний легких. Набор данных содержит большой набор рентгенограмм и клинических данных об около 65 000 случаев госпитализации. Информационный массив содержит сводную информацию о госпитализациях пациентов, установленных заболеваниях, назначенных курсах лечения и результатах терапевтического воздействия. Рентгенографические исследования, выполненные в рамках одной или нескольких госпитализаций, вместе с параметрами гемограммы, биологической плазмы и тканевой жидкости, включены в категории лабораторной биохимической аналитики набора данных. Всего имеется 1 630 категорий различных лабораторных испытаний. Для экспедитивного изучения нами была взята небольшая часть набора данных из 144 пациентов с диагностированными легочными заболеваниями без сопутствующих заболеваний других категорий и 145 пациентов контрольной группы, у которых были ди-

агностированы заболевания, не связанные с легочными патологиями, а именно нарушения сна. Каждый пациент был помечен, соответственно, как больной или здоровый. Среди более чем 18 000 пациентов с заболеваниями легких по кодам ICD-9 480-508 были отобраны 144 пациента, у которых не были диагностированы другие сопутствующие заболевания, поскольку последние могли повлиять на назначения, курс лечения и рентгеновские снимки. Поскольку каждый пациент, страдавший заболеваниями легких, скорее всего, проходил рентгенологическое обследование несколько раз, нами были рассмотрены только первые снимки, которые были сделаны при поступлении пациента в больницу. Изучение только первых снимков является целесообразным, поскольку во время лечения состояние легких меняется, т.к. они поддавались воздействию назначенных лекарств, а картина симптомов часто становится нечеткой или полностью исчезает. Кроме того, снимки были сделаны на разном оборудовании, в разном качестве, с разным разрешением и в разных плоскостях. Проблемы обработки качества изображений решены нами на этапе предварительной обработки данных, и при первоначальном отборе пациентов были взяты только пациенты с фронтальным положением грудной клетки во время рентгенологического обследования. Пациент мог находиться в положении стоя или лежа, но его грудная клетка была повернута непосредственно к сканеру. У некоторых пациентов было диагностировано несколько заболеваний легких одновременно. Среди пациентов без диагностированных заболеваний легких было отобрано 145 пациентов с заболеваниями, которые с наименьшей степенью вероятности повлияли на типичные изменения медицинских показателей пациентов с заболеваниями легких. Нами рассмотрены пациенты с единичными заболеваниями из группы 327.00–327.59, согласно коду ICD-9-CM, в основном это бессонница и расстройства сна. После отбора пациентов в соответствии с вышеуказанными критериями был проведен анализ клинических показателей. В соответствии с диагностированным заболеванием пациенты проходили различные клинические обследования. Существовало не менее 200 записей результатов по клиническим показателям для каждого пациента. При этом ряд показателей измерялся несколько раз, как в разные периоды госпитализации, так и во время одной госпитализации, чтобы отобразить динамику изменения значений. В связи с этим нами оставлены только те результаты клинических испытаний, которые соответствовали моменту их первого определения. В среднем на каждого пациента приходилось около 65 лабораторных исследований. Сами клинические испытания варьировались от пациента к пациенту в зависимости от особенностей заболеваний и назначенного лечения. Среди других клинических данных в исходном наборе данных есть информация о названиях лекарств и их дозах, назначаемых пациентам. Поскольку они не влияют на процесс бинарной классификации, а в большей степени касаются прогнозирования течения заболеваний, в настоящем исследовании эти данные не были приняты во внимание. Наиболее распространенными диагнозами заболеваний легких, выявленными у более чем 20 пациентов из набора данных, являются пневмония, в том числе инфекционная, хроническая обструкция дыхательных путей, обструктивный хронический бронхит и астма. Среди наиболее распространенных лабораторных тестов, назначенных и проведенных более чем 60 пациентам из набора данных при их первом поступлении, были следующие анализы: гемоглобин, креатинин, гематокрит, уровень глюкозы в крови, глюкоза и билирубин в моче, аланинаминотрансфераза (АЛТ) и аспаратаминотрансфераза (АСТ).

3. Методы

3.1. Предварительная обработка изображений

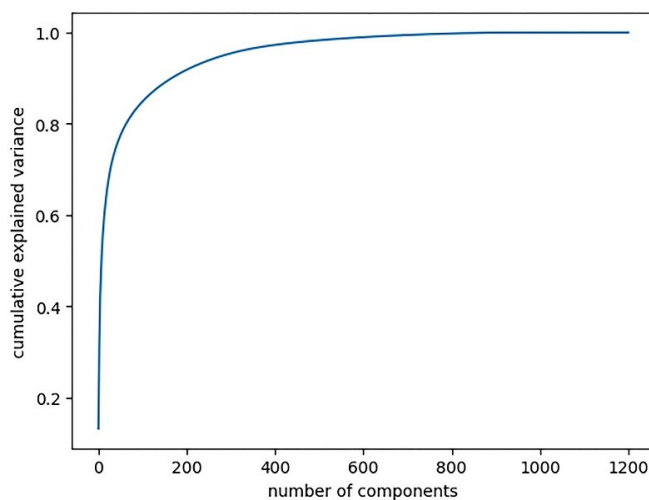


Рис. 1. Объясненное отношение дисперсии к количеству компонент

Поскольку рентгеновские снимки были выполнены на различном оборудовании, необходимо было преобразовать изображения к единому формату. Все изображения приведены нами к разрешению 64×64 и применена аугментация с помощью следующих методов: горизонтальный поворот, добавление шума (коэффициент 0.2) и преобразование изображений в монохромное без оттенков серого. Применение других методов аугментации представляется нецелесообразными. Большое число методов аугментации, как правило, применяется при отсутствии возможности существенно увеличить мощность набора данных. Мы осознанно сократили имеющийся большой набор данных до небольшого, чтобы показать преимущества многомодального глубокого обучения. При необходимости можно полностью отказаться от аугментации как метода предобработки изображений. На следующем этапе применяется мэппинг для выбора наиболее значимых признаков изображения с помощью методов главных компонент (PCA, Principal Component Analysis) и рекурсивного подавления признаков (RFE, Recursive Feature Elimination). Рисунок 1 отражает отношение выбранных главных компонент изображения к кумулятивному размеру объясненной дисперсии данных. Алгоритм определил 117 определяющих признаков, которые в дальнейшем использовались как порог при запуске метода RFE.

3.2. Предварительная обработка клинических данных

В использованном наборе данных представлены записи о разнообразных лабораторных тестах, назначенных для пациентов с различными заболеваниями. Необходимо было выделить группу лабораторных тестов, которые были назначены и проведены как пациентам с заболеваниями легких, так и пациентам с расстройствами сна. Из 1 630 значений всех возможных медицинских анализов и исследований необходимо было взять приемлемое меньшее количество для решения задачи бинарной классификации. Визуализация значений, выходящих за пределы нормальных диапазонов, представлена на рис. 2 желтым цветом. В верхней строке представлены данные для здоровых пациентов, в нижней строке — для пациентов с легочными заболеваниями.

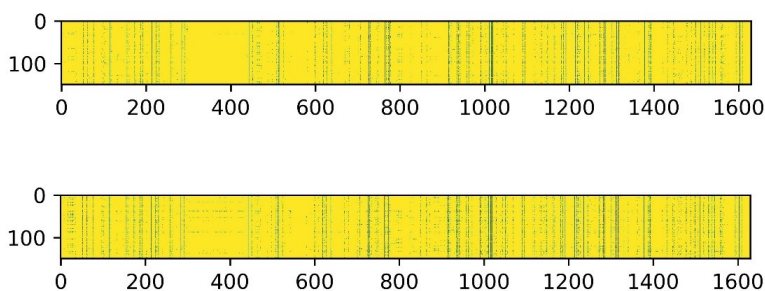


Рис. 2. Клинические анализы: нормальный (желтый) и патологический (зеленый)

Все лабораторные исследования используемого набора данных и, соответственно, их показатели, разделены на четыре группы: 1) назначенные практически всем пациентам (стандартные общие анализы крови и др.); 2) являющиеся специфическими для пациентов с подозрением на заболевание легких; 3) являющиеся специфическими для пациентов с подозрением на заболевания, приводящими к нарушениям сна; 4) те, которые встречаются у пациентов обеих групп эпизодически, всего в одном или нескольких случаях. Соответственно, исследования последней группы были исключены из набора данных. Была произведена операция уплотнения матрицы значений лабораторных исследований и количество параметров снижено до 521.

Для того чтобы расширить набор данных, нами выполнена генерация искусственных образцов (sample) лабораторных исследований. Такой шаг является часто распространенным приемом обработки медицинских клинических данных в связи со сложностью сбора данных. Обычно, для такой генерации табличных данных используются разнообразные методы, например, зашумляющий автоэнкодер, генеративные состязательные сети с использованием метрики Вассерштайна (WGAN, Wasserstein Generative Adversarial Nets), портал MDCClone.com и др. В будущих исследованиях мы планируем вернуться к вопросу обработки исходного набора данных из 6 500 записей о пациентах для проверки работы модели на больших данных. В данном исследовании мы решили сделать ограничение по количеству реальных данных и сгенерировать такой же объем данных, какой был сгенерирован для рентгеновских снимков, — 1 156. Нами использован автоэнкодер, который зашумлял исходные данные о значениях клинических исследований пациентов, а затем восстанавливал их, получая похожие результаты. Соответствие между сгенерированными данными клинических испытаний и синтетическими рентгенограммами устанавливалось согласно следующему правилу.

Пусть исходный набор данных описывается следующим образом: $P = \{CXR; EMR\}$, где CXR — рентгенологический снимок грудной клетки; EMR — табличные данные о результатах лабораторных исследований пациента. Тогда каждый пациент описывается как $P_i = CXR_i; EMR_i$, где $i \in [1..289]$. Набор сгенерированных с помощью аугментации рентгенологических снимков может быть описан как

$$CXR_i = \{CXR_{i1}; CXR_{i2}; CXR_{i3}\}. \quad (1)$$

Набор синтезированных данных о лабораторных исследованиях может быть представлен как

$$EMR_i = \{EMR_{i1}; EMR_{i2}; EMR_{i3}\}. \tag{2}$$

Новый образец (пациент) объединяет информацию о снимке и анализах следующим образом:

$$P_{ij} = \{CXR_{ij}; EMR_{ij}\}, \tag{3}$$

где $i \in [1..3]$, поскольку применены три метода аугментации и выполнены три запуска зашумляющего автоэнкодера. В результате набор синтезированных пациентов может быть описан как

$$P_j = \{P_{j1}; P_{j2}; P_{j3}\}. \tag{4}$$

Наконец, расширенный набор данных включил в себя исходных и сгенерированных пациентов: $N = P \cup P_j$. Такой прием может применяться к набору данных малой мощности в любой предметной области, где имеется необходимость в обработке разных модальностей.

3.3. Мультимодальная модель позднего слияния

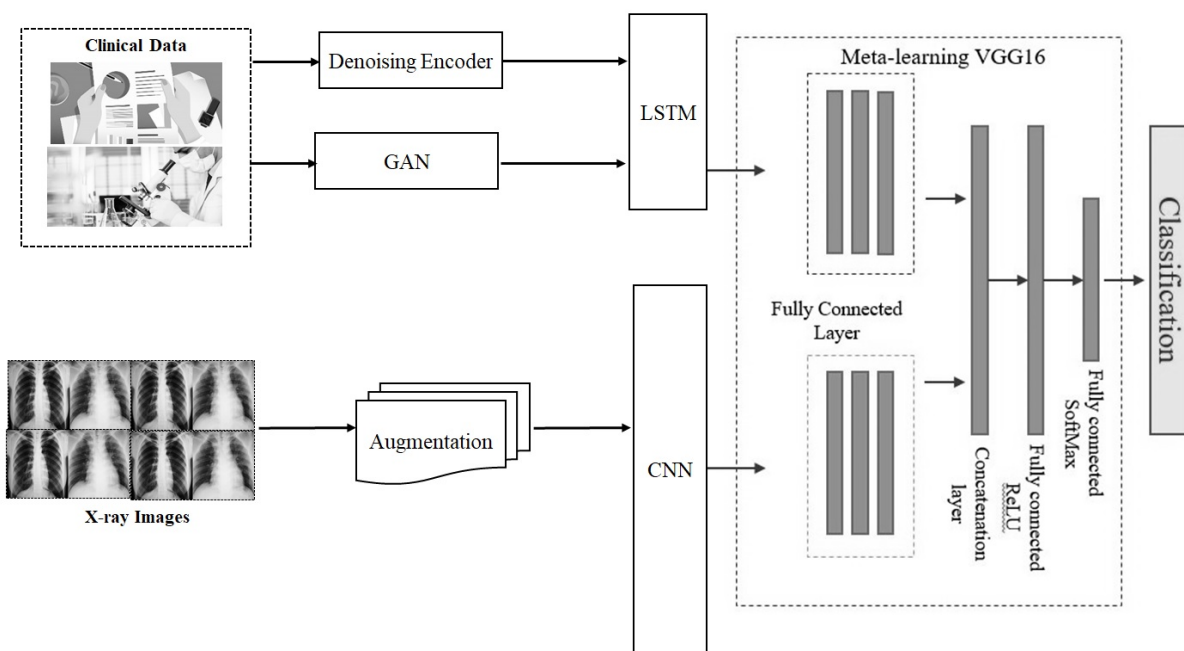


Рис. 3. Модель позднего слияния мультимодальных данных для классификации заболеваний легких

Общая схема мультимодальной модели позднего слияния представлена на рис. 3. На первом этапе исследования применены различные методы предобработки графических изображений: метод рекурсивного подавления признаков (RFE), мэширующий метод главных компонент (PCA) с порогом 0.5 и их комбинация. Изображения, полученные после применения первого из перечисленных методов, оказались наиболее удачным выбором для последующего применения в обучении нейронной сети. Нами использовался метод опорных векторов (SVM) для выравнивания объектов различных модальностей. Завершение обучения по объединенной карте признаков проводилось с учетом соответствия клинических данных рентгеновским изображениям. Для каждой комбинации параметров модель

запускалась не менее 5 раз, чтобы получить средние результаты в условиях случайного выборочного разделения данных на обучающие (60%), валидационные (20%) и тестовые (20%) наборы. Обучение включало раннюю остановку, чтобы предотвратить переобучение. Модель запускалась с различным размером пакета (batch size): от 2 до 10. Меньший размер пакета давал меньший эффект флуктуации точности в конце эпох обучения. Для максимального нивелирования этого эффекта была использована формула (5) адаптивного размера пакета, предложенная нами в работе [15].

$$b_i = b_{i-1} + \frac{N_{i-1} \bmod b_i}{\log N_{i-1}}, \quad (5)$$

где b_i — размер пакета в текущую эпоху обучения, b_{i-1} — размер пакета в предыдущую эпоху обучения, N_{i-1} — размер набора данных в предыдущую эпоху обучения. Процесс обучения проводился на 50, 100 и 200 эпохах с ранней остановкой при наличии признаков переобучения. Модель Attention учитывала два последних временных шага в качестве входных характеристик, помогающих понять недавнюю тенденцию в обучении. Для количественной оценки разницы между прогнозируемыми и фактическими значениями использовалась функция потерь со среднеквадратичной ошибкой. Оптимизатор Adam был использован для эффективного обновления весов модели. В качестве параметра градиентного спуска была выбрана скорость обучения, равная 0.01, что обеспечивало превалирование лучшей точности по сравнению со скоростью работы модели. Оценка точности проводилась неоднократно. Эксперименты проводились не менее пяти раз с контролем доверительного интервала в 10% при разделении набора данных на обучающую, валидационную и тестовую выборки. Параметры контролировались во время новых обучающих запусков, с различными значениями генератора случайных состояний, различными значениями пакета и различным количеством эпох.

3.4. Метрики, используемые при оценке работы модели

Для расчета точности используется следующая формула:

$$Accuracy = \frac{\text{True Sick Patients} + \text{True Healthy Patients}}{\text{Number of Patients}}, \quad (6)$$

где True Sick Patients — правильно классифицированные пациенты, имеющие заболевания легких; True Healthy Patients — правильно классифицированные пациенты, не имеющие заболеваний легких; Number of Patients — общее количество пациентов в наборе данных. Модель оптимизировала целевую функцию точности классификации на основе достижения наименьшей среднеквадратичной ошибки в подмножестве валидационных обучающих данных:

$$MSE = \frac{1}{N_{\text{val}}} \sum (y_i - y_{i\text{predicted}})^2, \quad (7)$$

где y_i — i -й пациент из валидационной подвыборки набора данных; $y_{i\text{predicted}}$ — предсказание модели классификации для i -го пациента из валидационной подвыборки набора данных; N_{val} — количество пациентов в валидационной подвыборке набора данных.

4. Результаты и обсуждение

В этом разделе исследована производительность предлагаемой модели при различных параметрах запуска. Сравнение результатов обучения с помощью моделей DenseNet121,

DenseNet169 и ResNet50 на нашем наборе данных показало, что модель DenseNet169 дает наилучшую точность. Модель Attention, по сути являющаяся трансформером с реализованным механизмом самовнимания, предоставляемым фреймворком Keras, показала наилучшие результаты для лабораторных данных. Используя показатель точности, мы сравнили модели с одной модальностью и предложенную модель позднего слияния при обучении на многомодальном наборе данных, который включает информацию об исходных 289 пациентах, а также набор данных с включенными сгенерированными клиническими данными и дополненными рентгеновскими изображениями. Неизменными условиями экспериментов были следующие: фиксированный размер партии, метод генерации искусственных клинических данных — автоэнкодер. Данные представлены в таблице ниже в порядке возрастания точности. В исследуемой задаче классы сбалансированы, поэтому можно не использовать большое количество метрик. Обучение многомодальной модели запускалось не менее пяти раз. В экспериментах с адаптивным размером пакета мы констатировали одновременное повышение точности и увеличение потерь. Это могло означать, что наша модель становилась лучше с точки зрения точности при любом установленном нами пороговом значении. Этот эффект станет предметом дальнейших исследований. В целом, все методы показывают разумную точность. Однако использование многомодальной модели повышает точность классификации как минимум на 2.89% по сравнению с лучшим результатом одномодальной модели на таких же данных.

Таблица 1. Сравнение точности одномодальных и многомодальных методов классификации

Метод	Точность	Потери
Одна модальность, снимки, 289 пациентов	87.45	4.47
Одна модальность, снимки, 289 пациентов	85.79	5.36
Многомодальная модель позднего слияния, 289 пациентов	90.34	2.17
Одна модальность, снимки, 1156 пациентов	86.11	3.70
Одна модальность, лабораторные анализы, 1 156 пациентов	88.15	4.06
Многомодальная модель позднего слияния, 1 156 пациентов	91.30	1.52

Заключение

В данном исследовании предложена архитектура многомодальной модели классификации легочных заболеваний, основанная на обработке рентгеновских снимков грудной клетки и клинических данных лабораторных исследований больных. Предлагаемая модель является альтернативой стандартному подходу обработки данных одиночной модальности. В ходе исследования имеющийся набор данных был дополнен искусственно сгенерированными данными с рентгенологическими снимками пациентов, зашумленных автоэнкодером, и искусственно сгенерированными табличными данными о результатах медицинских лабораторных анализов. Также предложен алгоритм установления соответствия между искусственно сгенерированными данными различной модальности. Выполнено обучение нескольких моделей на обеих модальностях и выбрана предобученная модель DenseNet169 для обработки рентгенологических снимков легких и модель Attention, являющаяся трансформером, для обработки двумерных данных. После обучения выбранных моделей на отдельных модально-

стях использован метод позднего слияния для формирования метамодели мультимодальных данных. Для установления эффективности предложенной мультимодальной модели проведена серия экспериментов на исходном и искусственно расширенном наборах данных, которые показали, что модель имеет более высокую точность классификации в сравнении с традиционными методами.

Исследование выполнено при поддержке регионального гранта Российского научного фонда № 23-21-10009.

Литература

1. 2019 I.-1.V. Chapter X. Diseases of the respiratory system (J00-J99). 2019. URL: <https://icd.who.int/browse10/2019/en> (дата обращения: 27.10.2019).
2. 1998 I.-9.V. Diseases of the respiratory system (460-519) Chapter Pneumonia and influenza (480-488). 2019. URL: https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-services-plan/diag-codes_respiratory.pdf (дата обращения: 27.10.2019).
3. Sen I., Hossain M.I., Shakib M.F.H., *et al.* In Depth Analysis of Lung Disease Prediction Using Machine Learning Algorithms // Communications in Computer and Information Science. 2020. Vol. 1241. DOI: 10.1007/978-981-15-6318-8_18.
4. Bharati S., Podder P., Mondal M.R.H. Hybrid deep learning for detecting lung diseases from X-ray images // Informatics in Medicine Unlocked. 2020. Vol. 20, no. 100391. DOI: 10.1016/j.imu.2020.100391.
5. Mustafa E., Selim A. Detection of lung disorders using embedded and wrapper feature selection methods // Kahramanmaraş Sutcu Imam Universitesi Muhendislik Bilimleri Dergisi. 2022. Vol. 25, no. 100391. P. 452–460. DOI: 10.17780/ksujes.1138377.
6. Gu Y., Lu X., Yang L., *et al.* Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs // Computers in Biology and Medicine. 2018. Vol. 103. P. 220–231. DOI: 10.1016/j.compbiomed.2018.10.011.
7. Setio A.A.A., Traverso A., de Bel T., *et al.* Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge // Medical Image Analysis. 2017. Vol. 42. P. 1–13. DOI: 10.1016/j.media.2017.06.015.
8. Zhu W., Liu C., Fan W., Xie X. DeepLung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification // 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018, Lake Tahoe, NV, USA, March 12-15, 2018. IEEE Computer Society, 2018. P. 673–681. DOI: 10.1109/WACV.2018.00079.
9. Kong W., Hong J., Jia M., *et al.* YOLOv3-DPFIN: A Dual-Path Feature Fusion Neural Network for Robust Real-Time Sonar Target Detection // IEEE Sensors Journal. 2020. Vol. 20, no. 7. P. 3745–3756. DOI: 10.1109/JSEN.2019.2960796.

10. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 - 18th International Conference Munich, Germany, October 5 - 9, 2015, Proceedings, Part III. Vol. 9351 / ed. by N. Navab, J. Hornegger, W.M.W. III, A.F. Frangi. Springer, 2015. P. 234–241. Lecture Notes in Computer Science. DOI: 10.1007/978-3-319-24574-4_28.
11. Kallianos K., Mongan J., Antani S., *et al.* How far have we come? Artificial intelligence for chest radiograph interpretation // Clinical Radiology. 2019. Vol. 74, no. 5. P. 338–345. DOI: 10.1016/j.crad.2018.12.015.
12. Bhandary A., Prabhu G.A., Rajinikanth V., *et al.* Deep-learning framework to detect lung abnormality – A study with chest X-Ray and lung CT scan images // Pattern Recognition Letters. 2020. Vol. 129. P. 271–278. DOI: 10.1016/j.patrec.2019.11.013.
13. Bharati S., Podder P., Paul P.K. Lung Cancer Recognition and Prediction According to Random Forest Ensemble and RUSBoost Algorithm Using LIDC Data // Int. J. Hybrid Intell. Syst. 2019. Vol. 15, no. 2. P. 91–100. DOI: 10.3233/HIS-190263.
14. Behzadi-khormouji H., Rostami H., Salehi S., *et al.* Deep learning, reusable and problem-based architectures for detection of consolidation on chest X-ray images // Computer Methods and Programs in Biomedicine. 2020. Vol. 185. P. 105162. DOI: 10.1016/j.cmpb.2019.105162.
15. Kumar S., Ivanova O., Melyokhin A., Tiwari P. Deep-learning-enabled multimodal data fusion for lung disease classification // Informatics in Medicine Unlocked. 2023. Vol. 42. P. 101367. DOI: 10.1016/j.imu.2023.101367.

Иванова Ольга Николаевна, к.п.н., доцент, кафедра системного программирования, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Кумар Сэчин, PhD, в.н.с., лаборатория больших данных и машинного обучения, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Цымблер Михаил Леонидович, д.ф.-м.н., доцент, кафедра системного программирования, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Иванова Елена Владимировна, к.ф.-м.н., доцент, кафедра системного программирования, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

CLASSIFICATION OF MULTIMODAL LUNG DISEASE DATA BASED ON LATE FUSION OF MODALITIES

© 2024 O.N. Ivanova, S. Kumar, M.L. Zymbler, E.V. Ivanova

South Ural State University (pr. Lenina 76, Chelyabinsk, 454080 Russia)

E-mail: onivanova@susu.ru, kumars@susu.ru, mzym@susu.ru, elena.ivanova@susu.ru

Received: 21.09.2023

With the development of technology, high-quality X-rays have become available for the diagnosis of lung diseases with the help of radiologists. However, the diagnostic process takes a lot of time and depends on the availability of specialists in a medical institution. At the same time, patient information may include not only chest X-rays of different quality, but also the results of medical tests, doctor's notes and prescriptions, information about taking medications, and others. In this study, we propose a model for the classification of pulmonary diseases based on multimodal data on clinical studies of patients and radiographic images. When preparing the data, we used various methods of generating artificial samples for both images and tabular data on the results of laboratory studies. We have proposed a method for establishing a correspondence for generated samples between modals. The proposed multimodal model has a late fusion architecture. We conducted experiments on datasets with one modality and two modalities. Our model showed accuracy 5.5% higher than models based on single-modality (91.3% vs. 86.11% on a dataset of 1,156 patients).

Keywords: multimodal data, lung diseases, deep learning, late fusion.

FOR CITATION

Ivanova O.N., Kumar S., Zymbler M.L., Ivanova E.V. Classification of Multimodal Lung Disease Data Based on Late Fusion of Modalities. Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering. 2024. Vol. 13, no. 1. P. 74–86. (in Russian) DOI: 10.14529/cmse240105.

This paper is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 License which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is properly cited.

References

1. 2019 I.-1.V. Chapter X. Diseases of the respiratory system (J00-J99). 2019. URL: <https://icd.who.int/browse10/2019/en> (accessed: 27.10.2019).
2. 1998 I.-9.V. Diseases of the respiratory system (460-519) Chapter Pneumonia and influenza (480–488). 2019. URL: https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-services-plan/diag_codes_respiratory.pdf (accessed:1998).
3. Sen I., Hossain M.I., Shakib M.F.H., *et al.* In Depth Analysis of Lung Disease Prediction Using Machine Learning Algorithms. Communications in Computer and Information Science. 2020. Vol. 1241. DOI: 10.1007/978-981-15-6318-8_18.
4. Bharati S., Podder P., Mondal M.R.H. Hybrid deep learning for detecting lung diseases from X-ray images. Informatics in Medicine Unlocked. 2020. Vol. 20, no. 100391. DOI: 10.1016/j.imu.2020.100391.
5. Mustafa E., Selim A. Detection of lung disorders using embedded and wrapper feature selection methods. Kahramanmaraş Sutcu Imam Universitesi Muhendislik Bilimleri Dergisi. 2022. Vol. 25, no. 100391. P. 452–460. DOI: 10.17780/ksujes.1138377.

6. Gu Y., Lu X., Yang L., *et al.* Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs. *Computers in Biology and Medicine*. 2018. Vol. 103. P. 220–231. DOI: 10.1016/j.compbimed.2018.10.011.
7. Setio A.A.A., Traverso A., de Bel T., *et al.* Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. *Medical Image Analysis*. 2017. Vol. 42. P. 1–13. DOI: 10.1016/j.media.2017.06.015.
8. Zhu W., Liu C., Fan W., Xie X. DeepLung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification. 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018, Lake Tahoe, NV, USA, March 12-15, 2018. IEEE Computer Society, 2018. P. 673–681. DOI: 10.1109/WACV.2018.00079.
9. Kong W., Hong J., Jia M., *et al.* YOLOv3-DPFIN: A Dual-Path Feature Fusion Neural Network for Robust Real-Time Sonar Target Detection. *IEEE Sensors Journal*. 2020. Vol. 20, no. 7. P. 3745–3756. DOI: 10.1109/JSEN.2019.2960796.
10. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 - 18th International Conference Munich, Germany, October 5 - 9, 2015, Proceedings, Part III*. Vol. 9351 / ed. by N. Navab, J. Hornegger, W.M.W. III, A.F. Frangi. Springer, 2015. P. 234–241. *Lecture Notes in Computer Science*. DOI: 10.1007/978-3-319-24574-4_28.
11. Kallianos K., Mongan J., Antani S., *et al.* How far have we come? Artificial intelligence for chest radiograph interpretation. *Clinical Radiology*. 2019. Vol. 74, no. 5. P. 338–345. DOI: 10.1016/j.crad.2018.12.015.
12. Bhandary A., Prabhu G.A., Rajinikanth V., *et al.* Deep-learning framework to detect lung abnormality – A study with chest X-Ray and lung CT scan images. *Pattern Recognition Letters*. 2020. Vol. 129. P. 271–278. DOI: 10.1016/j.patrec.2019.11.013.
13. Bharati S., Podder P., Paul P.K. Lung Cancer Recognition and Prediction According to Random Forest Ensemble and RUSBoost Algorithm Using LIDC Data. *Int. J. Hybrid Intell. Syst.* 2019. Vol. 15, no. 2. P. 91–100. DOI: 10.3233/HIS-190263.
14. Behzadi-khormouji H., Rostami H., Salehi S., *et al.* Deep learning, reusable and problem-based architectures for detection of consolidation on chest X-ray images. *Computer Methods and Programs in Biomedicine*. 2020. Vol. 185. P. 105162. DOI: 10.1016/j.cmpb.2019.105162.
15. Kumar S., Ivanova O., Melyokhin A., Tiwari P. Deep-learning-enabled multimodal data fusion for lung disease classification. *Informatics in Medicine Unlocked*. 2023. Vol. 42. P. 101367. DOI: 10.1016/j.imu.2023.101367.