DeepAAA: Clinically Applicable and Generalizable Detection of Abdominal Aortic Aneurysm Using Deep Learning

Клинически применимое и обобщаемое обнаружение аневризмы брюшной аорты с использованием глубокого обучения

**Абстракт**. Мы предлагаем методику глубокого обучения для выявления и количественного определения аневризм брюшной аорты (ААА). Состояние, которое приводит к более чем 10000 смертей в год в Соединенных Штатах, протекает бессимптомно, часто выявляется случайно и часто пропускается рентгенологами. Наша модельная архитектура представляет собой модифицированную трехмерную сеть U-Net в сочетании с подгонкой элипсов, которая выполняет сегментацию аорты и обнаружение AAA. В исследовании используются 321 КТ брюшной полости и таза, выполненные «Отделением радиологии больницы общего профиля штата Массачусетс» для обучения и валидации. Затем модель дополнительно тестируется на возможность обобщения на отдельном наборе из 57 обследований с различными демографическими характеристиками пациентов и характеристиками сбора данных, чем исходный набор данных. DeepAAA обеспечивает высокую производительность для обоих наборов данных (чувствительность / специфичность 0,91 / 0,95 и 0,85 / 1,0 соответственно), для контрастного и неконтрастного КТ-сканирования и работает с объемами изображений с различным числом изображений. Мы считаем, что DeepAAA превосходит литературные результаты радиологов по обнаружению случайных AAA. Ожидается, что модель может служить эффективным фоновым детектором в рутинных КТ-исследованиях для предотвращения пропуска случайных AAA.

**1. Введение**

Аневризма брюшной аорты (ААА), увеличение или расширение брюшной аорты, обычно встречается у мужчин старше 65 лет с распространенностью от 4 до 8% [5]. Необработанные аневризмы имеют тенденцию к росту и в конечном итоге могут разорваться с уровнем смертности, превышающим 90%. Поскольку большинство AAA являются бессимптомными до критического кровотечения, случайное обнаружение AAA становится критическим. Тем не менее, на рутинных обследованиях по компьютерной томографии (КТ) брюшной полости, только 65% AAA случайно идентифицированы [2]. Эта низкая частота сообщений затрудняет своевременное вмешательство пациентов. Обычно ААА диагностируют в первый раз, когда пациент уже подвергается риску разрыва [7]. Кроме того, в обычной клинической практике размер AAA определяется ручным измерением максимального диаметра аорты, что отнимает много времени и подвержено высокой вариабельности между замерами.

В течение последнего десятилетия было предложено множество методов компьютерной диагностики для автоматической сегментации аорты. Многие из этих предыдущих пособий использовали классические методы компьютерного зрения, которые требовали предварительных знаний, таких как внешние исходные точки для инициализации [3]. Вследствие постоянно растущей способности глубокого обучения нейронные сети в последнее время используются для сегментации аорты при КТ-ангиографии [6]. Тем не менее, эти предыдущие алгоритмы глубокого обучения фокусировались только на КТ-исследованиях с контрастом, в то время как случайная идентификация AAA при сканировании без контраста не менее важна, но более сложна. Кроме того, большинство предыдущих работ было посвящено задаче автоматической сегментации аорты [6,9,11], но существует очень мало исследований, в которых исследуется более прикладная задача обнаружения ААА, которая имеет гораздо большую клиническую значимость, чем просто выполнение только одной сегментации.

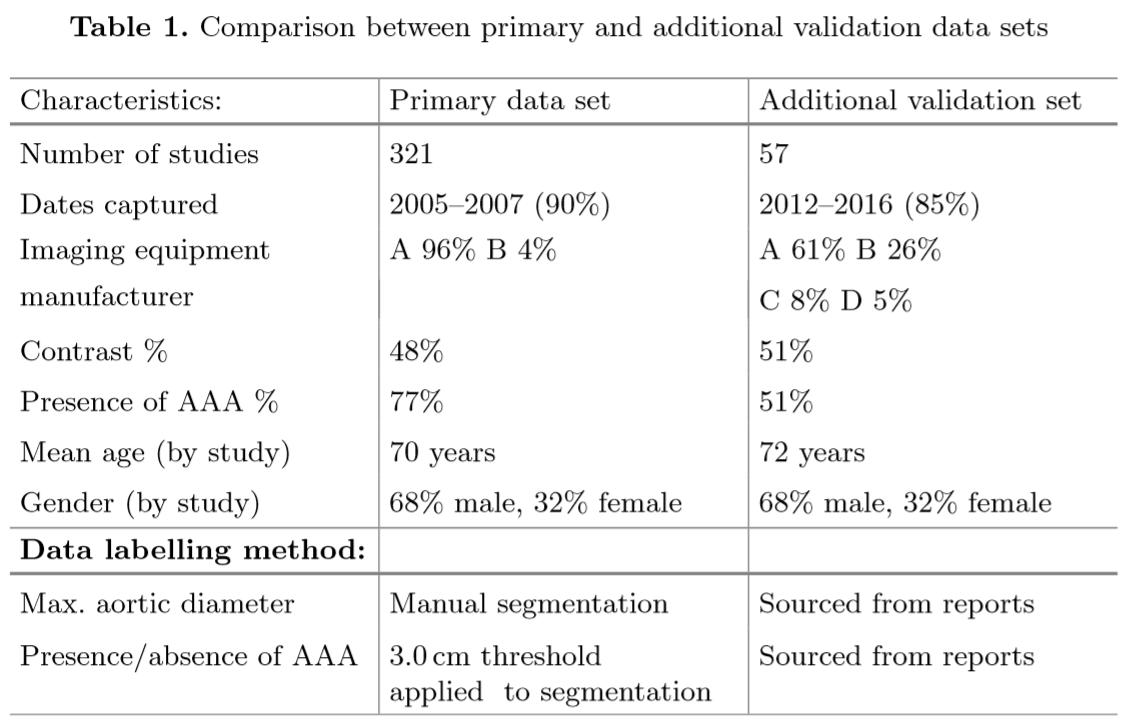
В этой статье мы демонстрируем решение для глубокого обучения (DeepAAA) для автоматической аортазегментации и АА-детекции в обеих контрастных и контрастных сериях КТ. В частности, мы разработали вариант трехмерной U-Net [1] для сегментации аорты при КТ брюшной полости. Предлагаемый способ обрабатывает серии с различным количеством изображений. Затем мы применяем подгонку эллипсом к сегментированным контурам аорты и оцениваем наибольший диаметр аорты. DeepAAA - это общее решение, обеспечивающее высокую частоту обнаружения AAA как при контрастном, так и неконтрастном КТ-сканировании и работающее с переменным разрешением изображения и толщиной среза. Кроме того, наше решение демонстрирует сильную обобщаемость и производительность относительно литературных значений чувствительности рентгенолога при обнаружении ААА.

**2**

Данные изображений состояли из контрастных и неконтрастных КТ-обследований живота и таза, выполненных в период с января 2005 года по апрель 2017 года Отделением радиологии больницы общего профиля штата Массачусетс. Исследователи получили одобрение местного Институционального контрольного совета на проект и отобрали два набора данных из базы данных. Эти два набора данных различаются по срокам их съемки и оборудованию для обработки изображений, как указано в Таблице 1.

**2.1 Первичный набор данных**

Первичный набор данных использовался для обучения и первоначальной проверки модели и содержал 321 исследование (223 уникальных пациента). Они были выбраны на основе поиска ключевых слов из отчета о положительных и отрицательных случаях ААА. Этот запрос был смещен в основном на исследования, сделанные в период между 2005 и 2007 годами. Из отобранных исследований было 217 (67,6%) мужчин и 104 (32,4%) женщины со средним возрастом 70,3 года; 153 (47,7%) КТ с контрастом и 168 (52,3%) без; 247 (76,9%) исследований с наличием ААА и 74 (23,1%) без ААА. Для каждого исследования осевой ряд использовался для сегментации аорты и обнаружения AAA. Толщина срезов изображений варьировалась от 2 до 10 мм, а количество изображений для каждой серии варьировалось от 40 до 384.



Поскольку многие замеры были сделаны несколькими людьми, была возможна оценка вариабельности между замерами. Из 153 контрастных исследований 124 были аннотированы как минимум двумя независимыми технологами, что привело к 517 парным сравнениям. Неконтрастные данные, однако, содержали только 10 исследований, в которых было выполнено более одной сегментации, в результате чего было получено только 16 парных сравнений. Среднее значение коэффициента Dice (Серестена) для оценки совпадения разных замеряльщиков в контрастных сериях составляло 0,95 ± 0,03, в то время как в неконтрастных сериях оно составляло 0,90 ± 0,08. Учитывая небольшое количество выборок, межпопулярная изменчивость неконтрастных данных не должна считаться определяющей, но предполагает примерно одинаковые уровни согласия. Для последующего анализа одна эталонная сегментация на набор данных была выбрана случайным образом в качестве основной истины.

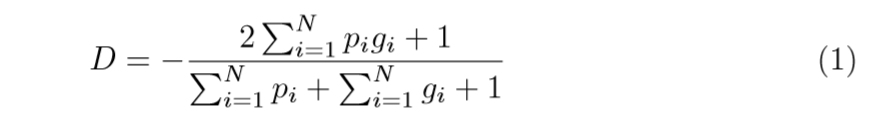
**2.2 Дополнительный набор для валидации.**

Для дополнительной валидации проводились проверки устойчивости модели к изменениям в оборудовании для визуализации, протоколах захвата отделения визуализации и демографии пациентов. Все эти факторы могут значительно меняться с течением времени в одном месте, и, таким образом, мы выбрали 57 исследований (57 уникальных пациентов), преимущественно захваченных между 2012 и 2016 годами для этого набора данных. Исследования были отобраны, чтобы включить смесь положительных и отрицательных случаев ААА посредством поиска по ключевым словам отчетов об исследованиях. Все отрицательные исследования были проверены вручную на отсутствие ААА. Для оценки модели относительно основанной на радиологии достоверности и проверки этапов постобработки, которые генерируют измерение AAA, максимальный диаметр аорты и наличие AAA были получены из радиологических отчетов, а не получены из ручной сегментации (как это было сделано для первичных данных устанавливать).

**3 метод**

Мы обнаруживаем AAA с помощью двух последовательных этапов: (1) сегментация аорты (2) контурирование аорты для оценки наибольшего диаметра поперечного сечения. Для сегментации брюшной аорты мы разработали свой вариант 3D U-Net [1], который принимает серии с различным количеством изображений. Как обсуждалось в разделе 2, наш набор данных содержал разные изображения и разные толщины срезов, поскольку исследования брюшной полости могут охватывать другие области тела, включая таз или грудную клетку. Таким образом, важно разработать алгоритм, адаптирующийся к изменчивости вдоль осевого измерения. Используемая нами архитектура 3D U-Net содержала 4 модуля понижающей / повышающей дискретизации (плюс уровень сужения), 2 сверточных слоя на модуль и 32 начальных элемента в сети. Размер сверточного ядра составлял 3 × 3 × 3 как на пути понижающей, так и на повышающей дискретизации, в то время как объемные ядра 3D составляли 2 × 2 × 1 для сохранения количества изображений. Batch normalization применяли перед каждой активацией ReLU, а регуляризацию dropout применяли на слое с узким местом со скоростью dropout 0,2. На выходном слое применяли сверточный слой 1 × 1 × 1 с активацией softmax по двум классам (фон и аорта). и пороговое значение 0,5, чтобы создать бинарную маску аорты.

Модель обучалась с помощью оптимизатора RMSprop с коэффициентом обучения 0,0001. Веса, выбранные для оценки, были такими, которые минимизировали потери в наборе валидации, которые в целом не были весами последней эпохи. Функция потерь была сглаженным отрицательным коэффициентом Дайса(Серестена): похожим, но не идентичным тому, который использовался в [8]. Суммирование проводится по всем N вокселям в скане, pi - это предсказанная вероятность аорты, а gi - основная классификация истинности для вокселя i. Дополнительные числа «+1» в числителе и знаменателе позволяют избежать деления на ноль.



Чтобы создать общий AAA-детектор, который работал с контрастным и неконтрастным КТ-сканированием, мы смешали оба типа КТ-изображений для обучения модели. Все эксперименты были выполнены с использованием библиотеки глубокого обучения Keras с бэкэндом Tensor Flow на **NVIDIA DGX-1 Volta**.

После сегментации аорты мы применили подгонку эллипсом [4] изображение за изображением к контурам аорты. Таким образом, наибольшие диаметры аорты (d) были заданы длинной осью эллипса. Для областей, где аорта не параллельна осевому КТ-сканированию, была применена коррекция угла для получения истинного диаметра аорты, то есть dcosθ, где θ - угол между секущей плоскостью аорты и осевым сканированием. Основываясь на определении AAA, предсказанные положительные результаты были исследованиями, где самый большой диаметр сегмента аорты был больше чем 3 см. Затем мы сравнили предсказанные результаты с аннотациями наземной правды.