

Система детектирования и анализа объектов на КТ-снимках в урологии

М. А. Руденко¹, А. В. Руденко², М.А. Крапивина³,
В.С. Лисовский⁴,

ФГАУО ВО «Крымский федеральный университет
им.В.И. Вернадского»

¹rudenko.ma@cfuv.ru, ²rudenkoandre@mail.ru,

³miladakrapivina@gmail.com, ⁴vllisovsky@mail.ru

Аннотация. В статье приводится анализ возможностей методов искусственного интеллекта для решения задачи нахождения и оценки объектов на снимках компьютерной томографии и рентгенографии с помощью методов глубокого машинного обучения.

Ключевые слова: компьютерное зрение, искусственная нейронная сеть, глубокое обучение, распознавание объектов, сверточная нейронная сеть, детектирование объектов

I. ВВЕДЕНИЕ

Одним из самых перспективных направлений использования искусственного интеллекта (ИИ) является применение в медицине, причем технологии ИИ приносят не только теоретическую пользу в ряде различных научных исследований, но и практическую. Применение ИИ дает возможность разработки новых, более эффективных подходов к анализу многомерной информации, например, при распознавании медицинских изображений. Это имеет большое значение при заболеваниях, диагностика которых основывается на исследовании медицинских изображений. К данному типу заболеваний относится мочекаменная болезнь, при которой диагноз чаще всего ставится по результатам проведения исследований методами рентгенографии и компьютерной томографии.

В данной статье будет рассмотрено применение компьютерного зрения (КЗ) в урологии, а именно детектирование и распознавание объектов свёрточными нейронными сетями (СНС) на снимках компьютерной томографии для проведения диагностики и лечения мочекаменной болезни.

Просмотр изображений, поиск, локализация почечных камней и ручная разметка снимков КТ сами по себе являются довольно сложными и трудоемкими процедурами, требуют значительных временных затрат специалистов и часто осложняются неполнотой и неточностью исходных данных (наличие различных видов искажений в снимках). Использование методов КЗ для

решения этих задач имеет ряд преимуществ и, главным образом, позволит ускорить процесс анализа медицинских снимков, снизить вероятность и риск ошибочных выводов и оптимизировать нагрузку на медицинских сотрудников.

Целью исследования является создание автоматизированной системы поддержки принятия врачебных решений (СПВР) в урологии на основе применения технологий компьютерного зрения для анализа и оценки результатов КТ.

Процесс поддержки принятия решений предполагает разработку нейросетевых моделей для обнаружения заданного класса объектов, а также системы нечеткого вывода об уровне принадлежности объекта классу. При планировании хирургического вмешательства для обнаруженных конкрементов необходимо вычислить характеристики, такие как размер, объем, масса, распределение плотности. СПВР должна сформировать заключения в понятном для врача виде.

Главной задачей, которая определяет эффективность разрабатываемой СПВР, является создание нейронной сети распознавания заданных классов объектов. Точность и адекватность нейросетевой модели зависит, в первую очередь, от выбранной технологии и качества обучения.

Современное развитие технологий компьютерного зрения предоставляет разработчику целый ряд библиотек и решений для различных платформ и на разных языках программирования. Выбор конкретного решения зависит от вида и характеристик изображений и объектов, которые необходимо обнаружить. Сложным этапом является обучение нейронной сети, которое требует привлечения специалистов-медиков для разметки изображений для обучения нейронной сети, а также формирование вывода и заключения относительно объекта, диагноза и пути лечения.

II. ОПИСАНИЕ МЕТОДИКИ

Компьютерная томография - медицинское рентгенологическое исследование, основанное на послойном исследовании структуры внутренних органов и систем, позволяющее получить рентгеновское изображение внутренних органов [1, 4]. Рентгеновские

Финансирование исследования проводится ФГБУ «Фонд содействия развитию малых форм предприятий в научно-технической сфере», программа «Старт», конкурс «Старт-Искусственный интеллект-1» (I очередь), заявка С1ИИ-112266, договор № 27ГС1ИИС12-D7/71365

лучи проходят сквозь тело человека, а компьютерная система формирует изображение как поперечных слоев (срезов), так и продольных. Результаты этой процедуры сохраняются в виде DICOM файлов. DICOM (Digital Imaging and Communications in Medicine) - объектный файл с теговой организацией для представления кадра изображения (или серии кадров) и сопровождающей информации (в виде DICOM тегов), который имеет 4-х ступенчатую структуру: пациент (patient) → исследование (study) → серия (series) → изображение (кадр или серия кадров, image) [4].

Таким образом, в результате проведения исследования врач-специалист получает большой набор изображений срезов внутренних органов по одному пациенту. Далее, с помощью специализированного программного обеспечения он просматривает полученные срезы для установки точной локализации, определения формы и размеров конкрементов, производит разметку среза. Размеченный срез представляет из себя собственное изображение и текстовый документ, содержащий информацию о выделенных объектах на данном изображении: их класс, координаты на срезе и размеры. После разметки срезов специалист определяет параметры объектов - размеры и плотность почек и камней.

Очевидны пути оптимизации этого процесса с помощью нейронной сети. Она должна получать изображения из DICOM файлов, распознавать их, детектировать на изображении объекты заданных классов, оценивать результаты распознавания и рассчитывать параметры (размеры, плотность, координаты) заданных медицинских объектов для дальнейшей 3D – визуализации объектов с картированием по плотности найденных объектов, характеризующих выбранную патологию, с целью поддержки процессов планирования хирургических вмешательств в области урологии.

III. ПРИМЕНЕНИЕ СНС ДЛЯ АНАЛИЗА КТ

Поиск объектов на медицинских изображениях является одной из важнейших задач при дифференциальной диагностике патологий на КТ. Задача поиска объектов на КТ снимках обычно состоит из двух подзадач: первая из них относится к классу задач обнаружения (детектирования) объекта, принадлежащего одному из заданных классов на срезе; вторая – в оценке принадлежности найденного объекта одному из заданных классов.

До недавнего времени задачи детектирования решались с использованием различных алгоритмов обработки изображений, в том числе: алгоритмов адаптивной фильтрации и усиления, алгоритмов, основанных на использовании гистограмм градиентов, алгоритмов каскадных классификаторов на основе метода Виолы-Джонса, алгоритмы, основанные на методах контурного анализа и др. [3].

Решения, основанные на обычных нейронных сетях с предварительным выделением признаков классификации объектов, а также стандартных глубоких сверточных нейронных сетей, применяемых в режиме просмотра последовательно перебираемых в скользящем режиме

окон, в которых проводится анализ [2]. Следует отметить, что такой подход является весьма затратным в вычислительном отношении, несмотря на все преимущества, которые изначально предоставляют нейронные сети с точки зрения точности классификации.

В последнее время для решения данной задачи широкое распространение получили алгоритмы, основанные на применении региональных глубоких сверточных нейронных сетей (Regional Convolutional Neural Networks, R-CNN) и сетей детектирования YOLO (You Only Look Once), которые принципиально ориентированы на решение задачи поиска объектов с одновременной их классификацией. По сравнению с методами, представленными выше, эти алгоритмы принципиально адаптированы для решения задачи поиска объектов на изображениях. Исходная реализация R-CNN базируется на использовании специальных алгоритмов предобработки – алгоритмов region-proposal-function, обеспечивающих предложение так называемых областей внимания, в которых потенциально могут находиться интересующие нас объекты. Такой «специализированный» подход предлагает сократить вычислительные затраты, а также позволяют добиться минимального времени определения местоположения объекта и высокой точности его классификация. К настоящему моменту имеется большое количество вариантов реализации подобных алгоритмов, которые достигли хороших показателей по данным критериям [3, 6, 8].

Многие алгоритмы решения задачи детекции объектов на изображениях с помощью СНС заключается в разбиении изображения на квадратные области, затем классификация этих областей на наличие объекта и классификация самого объекта. Таким образом, изображение просматривается дважды (один раз для определения областей, где есть объект, второй — для классификации этого объекта.) Этот способ работает долго и требует больших затрат вычислительных мощностей.

YOLO же использует другой принцип. Исходное изображение сжимается таким образом, чтобы получить квадратную матрицу размером 13 на 13, в каждой клетке которой записана информация о наличии объекта и классе этого объекта на соответствующей части картинка. Таким образом, YOLO просматривает картинку за один проход, что существенно увеличивает скорость обработки изображения. Данные характеристики позволяют использовать YOLO для задач распознавания объектов в реальном времени и на портативных устройствах. Отличительной особенностью YOLO является подход к решению задачи детекции, который предполагает одноразовую проход CNN для прогнозирования положения объекта и классификации нескольких кандидатов. Таким образом достигается сквозное обнаружения всех классов объектов. По результатам исследований [5, 8] сети YOLO являются наиболее эффективными по скорости и точности из всех нейронных сетей для детектирования объектов.

Для решения задачи диагностики в СПВР была выбрана архитектура YOLO, которая дает лучшие

результаты в области детектирования объектов на изображениях.

В настоящее время последней версией СНС архитектуры YOLO является YOLOv5, реализованная на фреймворке PyTorch, скорость распознавания которой составляет 140 кадров в секунду. YOLOv5 включает в себя несколько моделей, которые отличаются по количеству слоев и размеру, скоростью и точностью детектирования: YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x [2, 8].

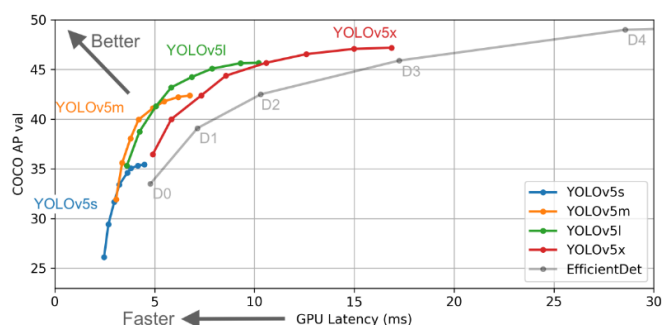


Рис. 1. Сравнение версий моделей YOLOv5

Использование сетей YOLO предполагает обучение на размеченных изображениях, которые представляют выборку заданных классов. Разметка изображений требует знаний специалиста в области хирургии и урологии. Это классифицируется как обучение с учителем, так как предполагается использование полного набора аннотированных данных (размеченного обучающего датасета) для тренировки системы ИИ на всех стадиях ее создания. Разработчиком YOLOv5 для получения хороших результатов обучения рекомендовано более 1500 изображений объектов одного класса, более 10000 размеченных экземпляров объектов для каждого класса [7].

В исследуемой области получить необходимое количество изображений для создания датасета нужного размера в медицине бывает проблематичным. Несмотря на то, что КТ в перспективе может давать множество материала для обучения (к примеру, толщина среза у современных 16-х срезовых аппаратов может достигать 0,75 – 0,6 мм, обучение проходит с использованием изображений, размеченных специалистами в предметной области, а это очень трудоемкий и долгий процесс.

Для дообучения нейросети и повышения точности детектирования можно использовать датасет, созданный на основе данных, которые были получены при первичном детектировании самой нейросетью и которые были проверены и откорректированы специалистом. Таким образом происходит оптимизация процесса дополнения обучающей выборки данными исследований новых пациентов.

На начальных этапах работы было использовано увеличение размера обучающего датасета путем клонирования исходных изображений с использованием фильтрации изображений – усилением резкости изображения (sharping), а также путем дублирования

файлов меток, содержащих разметки для изображений, на ближайшие соседние к ним.

Применение этих методов тиражирования изображений, хотя и ведут к снижению уникальности данных в датасете, оправданы на этапе создания обучающего датасета. В клинической практике с КТ сопутствуют различные искажения, такие как засвеченное изображение, наличие посторонних предметов, кроме того, вероятны проведения процедуры без контрастного вещества и предварительной подготовки пациента, что приводит к ухудшению четкости изображений. Что касается дублирование меток на соседние срезы – расстояние между ними, как правило, достаточно мало, чтобы это привело к возникновению существенной ошибки.

IV. МОДЕЛИ

Применяемая классификация объектов, которые могут быть детектированы нейросетью на изображениях КТ, содержит следующие классы:

- stone,
- left_kidney,
- right_kidney,
- left_kidney_pieloectasy,
- right_kidney_pieloectasy,
- staghorn_stones.

Для детектирования объектов была выбрана модель YOLOv5s, поскольку она содержит небольшое количество слоев, имеет достаточную скорость распознавания, файл весов имеет размер 14 МБ.

Первая версия модели YOLOv5s была обучена на датасете, содержащем 1586 изображений, в том числе 793 оригинальных изображений КТ, полученных из КТ у 21 пациента, 793 – обработанных изображений. Обучение проводилось 1200 эпох. Распределение изображений по классам неравномерно: классы left_kidney, stone, right_kidney значительно популярнее (Рис.2).

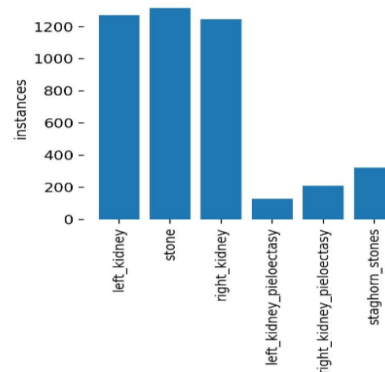


Рис. 2. Распределение классов первой версии обученной модели нейросети

Наблюдается прямая зависимость между количеством встреч объектов одного класса и точностью их распознавания. Тем не менее, достаточно высокая

точность распознавания наблюдаются для всех классов объектов.

Наибольшую точность модель показала при детектировании почек (рис.3). Интересно также то, что самый распространенный класс- «stones» имеет одну из самых низких точностей, что объясняется асимметрией обучающего датасета.

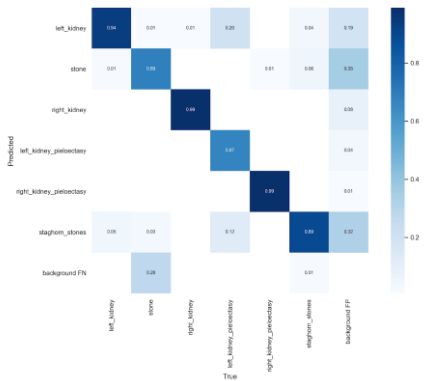


Рис. 3. Результаты обучения первой модели

Исходя из результатов первой модели можно сделать вывод, что для повышения точности распознавания камней необходимо учитывать их локализацию. Это можно сделать путем добавления дополнительных алгоритмов оценки результатов распознавания объектов. Авторами был разработан и внедрен в процесс детектирования нейросетью алгоритм оценки правдоподобия распознавания, который «отсекает» детектированные объекты «неправильной» локализации. В результате удалось добиться того, что объекты классов камней «stone» и «staghorn_stones» детектируются только в пределах объектов классов, принадлежащих почкам: правой и левой.

Вторая версия модели была обучена на датасете, содержащем 2075 изображений, в том числе 1582 оригинальных изображений, полученных из КТ 24 пациентов, остальные изображения были получены путем изменения оригиналов добавлением резкости. Обучение проводилось на 600 эпохах. Несмотря на увеличение количества снимков по сравнению с предыдущей моделью, соотношение классов не изменилось (рис. 4).

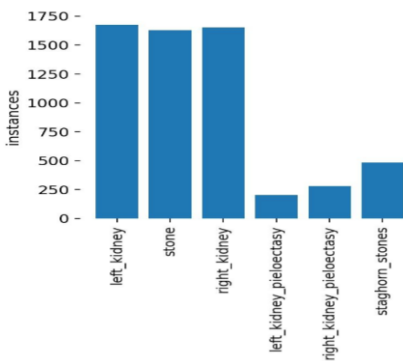


Рис. 4. Распределение классов второй версии обученной модели нейросети

Тем не менее очевиден прогресс и значительное увеличение точности работы нейросети (Рис 5).

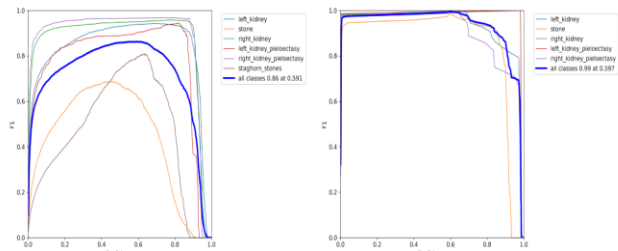


Рис. 5. Сравнение точностей первой и второй моделей

Вторая модель с высокой точностью определяет все заданные классы объектов (рис.6).

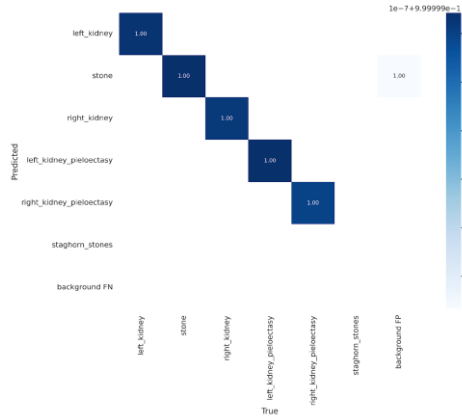


Рис. 6. Результаты обучения второй модели

Полученная модель нейронной сети для детектирования объектов на снимках КТ для проведения диагностик мочекаменной болезни была апробирована на базе Клинического медицинского многопрофильного центра имени Святителя Луки г. Симферополь. На рис. 7 приведен результат детектирования с помощью нейросети.

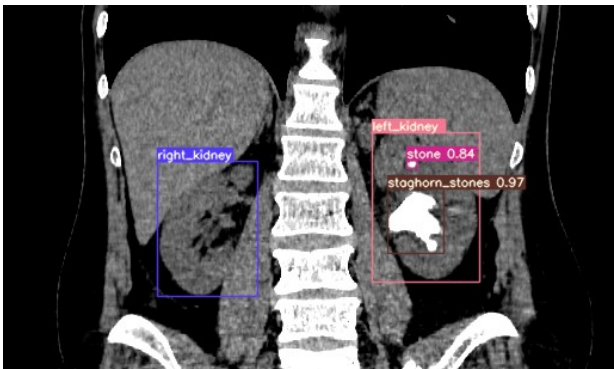


Рис. 7. Результат детектирования объектов на изображении КТ

Разработанная система детектирования и анализа объектов на КТ-снимках в урологии показала свою эффективность на объектах всех классов в процессе диагностирования и планировании хирургических вмешательств.

V. ЗАКЛЮЧЕНИЕ

Результаты исследования показали высокую эффективность применения искусственных нейронных сетей архитектуры YOLOv5 для детектирования заданных объектов на медицинских изображениях. Решение данных задач невозможно без широкого использования сверточных нейронных сетей. Однако наибольшую ценность представляют выводы, построенные на комплексной оценке результатов распознавания как всего изображения, так и отдельных объектов с помощью детектирования с использованием алгоритма оценки правдоподобия распознавания.

СПИСОК ЛИТЕРАТУРЫ

- [1] Борисов Д.Н., Кульнев С. В., Лемешкин Р. Н. ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ АНАЛИЗЕ ЦИФРОВЫХ ДИАГНОСТИЧЕСКИХ ИЗОБРАЖЕНИЙ// СОСТОЯНИЕ И ПЕРСПЕКТИВЫ РАЗВИТИЯ СОВРЕМЕННОЙ НАУКИ ПО НАПРАВЛЕНИЮ "ТЕХНИЧЕСКОЕ ЗРЕНИЕ И РАСПОЗНАВАНИЕ ОБРАЗОВ". 2019. С. 163-169 URL: <https://www.elibrary.ru/item.asp?id=41824272> (дата обращения: 28.04.2022).
- [2] Руденко М.А., Руденко А.В., Нечеткая модель классификации медицинских изображений на основе нейронных сетей, XXIV Международная конференция по мягким вычислениям и измерениям (SCM-2021). Сборник докладов. Санкт-Петербург. 26 – 28 мая 2021 г. СПб.: СПбГЭТУ «ЛЭТИ». 351 с. ISBN 978-5-7629-2864-92
- [3] Сирота, А. А., Митрофанова, Е. Ю., & Милованова, А. И. (2019). Анализ алгоритмов поиска объектов на изображениях с использованием различных модификаций сверточных нейронных сетей. Вестник ВГУ. Серия: Системный анализ и информационные технологии, (3), 123-137. <https://doi.org/10.17308/sait.2019.3/1313>
- [4] Хамад Ю.А., Симонов К.В. ПРОГРАММА ДЛЯ ОБРАБОТКИ И АНАЛИЗА ИЗОБРАЖЕНИЙ КОМПЬЮТЕРНОЙ ТОМОГРАФИИ И РЕНТГЕНОГРАФИИ ГРУДНОЙ КЛЕТКИ// ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ» (СФУ) 2020. URL: <https://www.elibrary.ru/item.asp?id=43965796> (дата обращения: 28.04.2022).
- [5] Priya, Dwivedi YOLOv5 compared to Faster RCNN. Who wins? / Dwivedi Priya. — Текст : электронный // towardsdatascience.com : [сайт]. — URL: <https://towardsdatascience.com/yolov5-compared-to-faster-rcnn-who-wins-a771cd6c9fb4> (дата обращения: 28.04.2022).
- [6] Srivastava, S., Divekar, A.V., Anilkumar, C. et al. Comparative analysis of deep learning image detection algorithms. J Big Data 8, 66 (2021). <https://doi.org/10.1186/s40537-021-00434-w>
- [7] Tips for Best Training Results. — Текст : электронный // docs.ultralytics.com : [сайт]. — URL: <https://docs.ultralytics.com/tutorials/training-tips-best-results/> (дата обращения: 28.04.2022).
- [8] YOLOv5: state-of-the-art модель для распознавания объектов. — Текст : электронный // neurohive.io : [сайт]. — URL: <https://neurohive.io/ru/papers/yolov5-state-of-the-art-model-dlya-raspoznavaniya-obektov/> (дата обращения: 28.04.2022).