

Построение системы поиска похожих опухолей с помощью глубоких сверточных сетей

Выполнил: Макаrchук Глеб

Научный руководитель: Беляев Михаил Геннадьевич

Московский физико-технический институт
(государственный университет)

Москва 2018

Содержание

- 1 Введение
- 2 Цель и задачи
- 3 Построение системы поиска
- 4 Анализ результатов

Введение

Решается задача построения системы поиска похожих опухолей головного мозга на трехмерных МРТ изображениях головы.

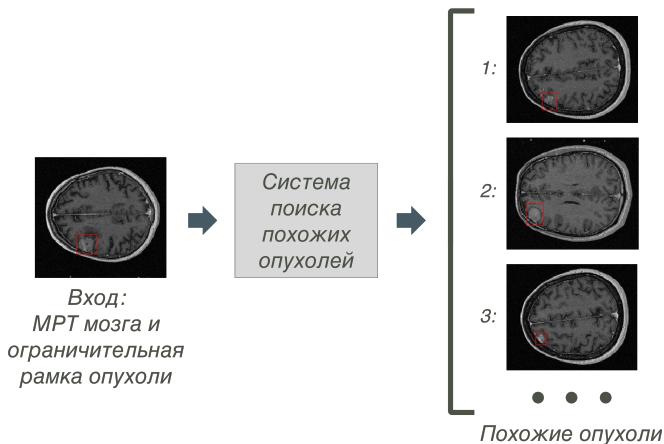


Рис. 1: Схема использования системы

Введение: размеченные МРТ данные в радиохирургии

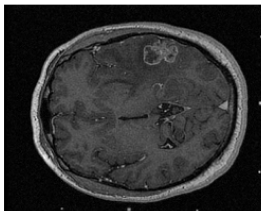


Рис. 2: Срез МРТ изображения мозга

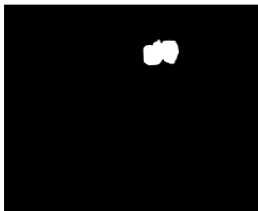


Рис. 3: Выделенная опухоль на срезе

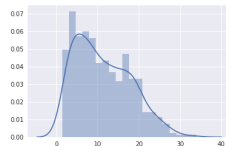


Рис. 4: Распределение линейных размеров опухолей

Имеется ≈ 1000 размеченных МРТ изображений содержащих всего ≈ 2400 опухолей. Размерность изображений: $200^3 \approx 10^7$. Известен тип опухолей (**3 класса**: метастазы 71%, менингиомы 17%, шванномы 12%), анатомическая область (**11 классов**), метка правого/левого расположения (**2 класса**).

Введение: мотивация и особенности задачи

Мотивация

Эффективная система поиска похожих опухолей позволит врачам повысить точность диагностики заболеваний, даст возможность прогнозировать развитие болезней и планировать лечение на основе похожих случаев.

Особенности

- Для опухолей имеется **различная разметка**: тип, контур, локализация.
- Задача **не формализована**. Сходство опухолей можно определять различными способами.
- Входные данные имеют **высокую размерность**.
- Размеченных данных относительно **мало**.

Цель и задачи исследования

Цель исследования

Построить систему поиска похожих опухолей, которая была бы полезна для клинического использования.

Задачи исследования

- Предложить метод построения системы поиска похожих опухолей.
- Оценить качество работы полученного алгоритма поиска, проанализировав сходство различных характеристик у опухолей при **ранжировании**.
- Получить оценку качества работы алгоритма от экспертов-врачей.

Обзор существующих подходов

- Использование вручную построенных признаков и обучение метрики сходства
- В качестве признакового описания используются представление с последних слоев нейронной сети:
 - Поиск похожих изображений глаз с наличием ретинопатии (Chung et al., 2017, MIT)
 - Поиск похожих MPT изображений изображений простаты (Shah et al., 2016, Siemens)

Особенности нашей задачи:

- Наличие различной разметки данных
- Отсутствие четких критериев сходства со стороны врачей-экспертов

Постановка задачи

X - множество трехмерных массивов, представляющих собой МРТ изображения. B - множество ограничительных рамок опухолей. Найти отображение в **пространство представлений**:

$$f_{repr} : X \times B \rightarrow \mathbb{R}^n$$

такое что что представления схожих опухолей **близки по l_1 , l_2 расстоянию**. Сходство определяется врачами-экспертами.

Оценка качества модели

Качество отображения f_{repr} будем оценивать с помощью:

- Свойства полезных признаков - возможности **предсказывать** характеристики опухолей по их представлениям:

$$f_K : \mathbb{R}^n \rightarrow K$$

где K - множество классов или \mathbb{R} .

- Свойства близости - n **ближайших соседей** опухолей по l_1 или l_2 имеют схожие характеристики.

Характеристики опухолей:

- Линейный размер опухоли
- Тип опухоли
- Анатомическая область опухоли
- Метка левого/правого расположения

Построение системы поиска

- 1 Обучим многоголовую сверточную нейронную сеть (CNN) на задаче **сегментации** опухолей, а также **классификации** опухолей по типу, анатомической области, левому-правому расположению.
- 2 Будем использовать обученную CNN для создания **представления опухоли** фиксированной размерности.
- 3 Поиск осуществляется с помощью **ранжирования опухолей** по l_1 , l_2 расстоянию в пространстве представлений.

Архитектура нейронной сети

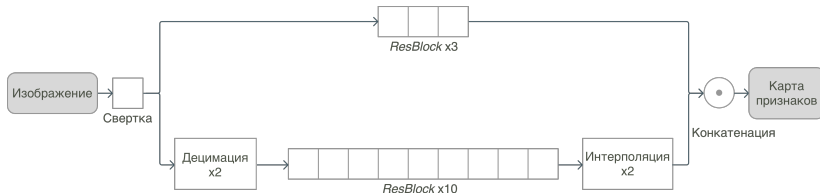


Рис. 5: Архитектура остова нейронной сети

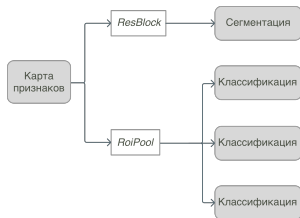


Рис. 6: Архитектура голов

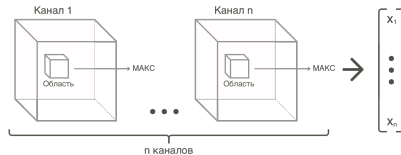


Рис. 7: Схема RoiPool

Архитектура нейронной сети

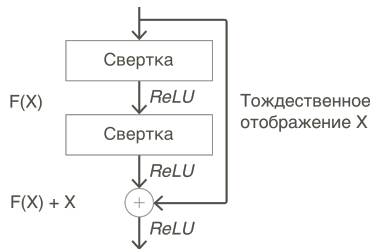


Рис. 8: Схема ResBlock

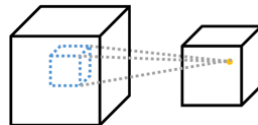


Рис. 9: Сверточный фильтр

- Верхний путь остова извлекает **локальные** признаки (малое количество сверток)
- Нижний путь остова извлекает **глобальные** признаки (большое количество сверток и децимация)

Обучение нейронной сети

- Обучение модели занимает **сутки**.
- Анализ проводился на **отложенных 20% изображений**.
Использовалось стратифицированное разбиение данных относительно типа опухолей.
- В процессе обучения использовались подобласти размера $120 \cdot 120 \cdot 120$ в силу технических ограничений.
- Сеть обучалась в течение 120 эпох. Шаг обучения уменьшался в 10 раз на 90 и 105 эпохах.
- В качестве функции потерь для задач сегментации использовалась бинарная кросс-энтропия, для задач классификации - мультиклассовая кросс-энтропия.

Свойство полезных признаков

Решается задача сегментации	Решаемые задачи классификации	Средняя абсолютная ошибка регрессии линейного размера	Точность классификации типа опухоли	Точность классификации анатомической области	Точность классификации левого-правого расположения
Да	Тип опухоли, анатомическая область, левое-правое расположение	1.64	0.93	0.58	0.81
Да	Тип опухоли, анатомическая область	1.92	0.91	0.54	0.57
Нет	Тип опухоли, анатомическая область, левое-правое расположение	2.4	0.92	0.63	0.57
Да	-	1.59	0.85	0.35	0.58
N/A	N/A	6.12	0.71	0.33	0.51

Таблица 1: Качество восстановления характеристик в зависимости от решаемых задач при обучении сети. Первые 2 столбца определяют на каких задачах была обучена нейронная сеть.

Визуализация представлений опухолей



Рис. 10: График линейных размеров опухолей после применения $2D$ PCA к 64-канальному представлению опухоли. Изменение цвета от синего к красному соответствует изменению размера от меньшего к большему.



Рис. 11: График типов опухолей после применения $2D$ PCA к 64-канальному представлению опухолей. Красный, зеленый и синий цвет означают метастазы, менингиомы и шванномы соответственно.

Зависимость от размерности представления

Размерность представления опухолей	Средняя абсолютная ошибка регрессии линейного размера	Точность классификации типа опухоли	Точность классификации анатомической области	Точность классификации левого-правого расположения
128	1.48	0.92	0.61	0.82
64	1.66	0.92	0.62	0.83
32	1.86	0.92	0.57	0.75

Таблица 2: Качество восстановления характеристик в зависимости от размерности представления опухолей.

Свойство близости

Количество ближайших опухолей	Относительное отличие линейных размеров			Сходство типов опухолей			Сходство анатомических областей		
	l1	l2	cos	l1	l2	cos	l1	l2	cos
2	0.217	0.229	0.273	0.881	0.874	0.889	0.311	0.301	0.298
3	0.227	0.24	0.288	0.874	0.864	0.88	0.312	0.296	0.286
5	0.238	0.253	0.312	0.87	0.863	0.876	0.289	0.278	0.284
10	0.267	0.281	0.345	0.854	0.858	0.869	0.268	0.271	0.268
20	0.299	0.314	0.394	0.835	0.833	0.846	0.249	0.249	0.257
2 случайным образом взятые опухоли	1.019			0.563			0.114		

Таблица 3: Сходство характеристик у ближайших опухолей по расстояниям l_1 , l_2 и косинусному сходству (\cos , формула 1). Сходство характеристик считается по формулам 2, 3.

$$\frac{A \cdot B}{|A|_2 \cdot |B|_2}$$

(1)

$$\frac{\sum_i^N |L_i - l_{ref}|}{N \cdot l_{ref}}$$

(2)

$$\frac{\sum_i^N [K_i == k_{ref}]}{N}$$

(3)

Пример поиска похожих опухолей

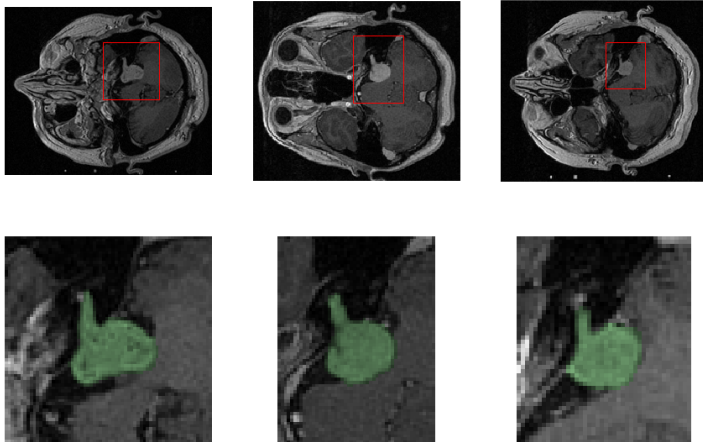


Рис. 12: Пример шванномы, и двух ближайших опухолей по l_2 расстоянию

Пример поиска похожих опухолей

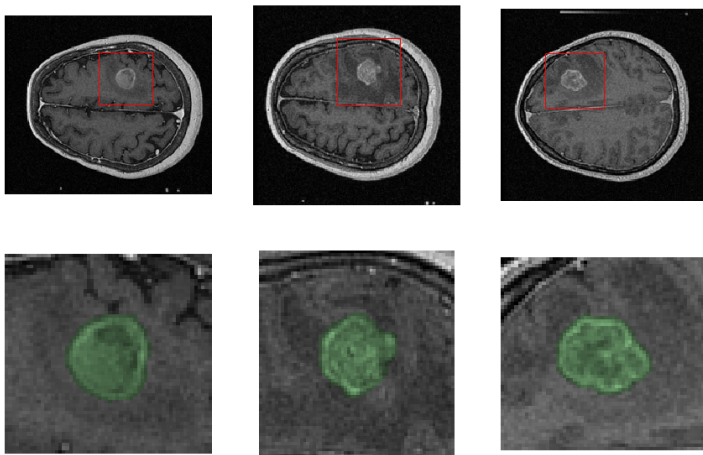


Рис. 13: Пример метастазы, и ее двух ближайших соседей

Отзыв врачей

Система была продемонстрирована врачам центра «Гамма-нож». Основные выводы из отзыва врачей после пользования системой:

- Опухоли в выдаче схожи по типу, размеру и локализации.
- Часто имелись **неформализованные** особенности. Похожие опухоли часто имели схожую форму или следы.
- Система особенно полезна для анализа сложных случаев нестандартных опухолей.
- Система может быть эффективно использована для обучения новых специалистов и использования в отделениях, где нет специалистов по опухолям мозга.
- Часть ошибок при поиске похожих опухолей происходила по причине неправильной разметки или некорректных данных.

Заключение

- Построена система поиска похожих опухолей.
- Проведена оценка качества поиска с помощью анализа характеристик опухолей в поисковой выдаче.
- Получено подтверждение врачей о полезности использования системы в реальных условиях.

Спасибо за внимание!
Вопросы?