

# Нейронные сети для анализа томограмм сетчатки глаза

В данной репозитории представлены модели свёрточных нейронных сетей, созданные для классификации томограмм сетчатки глаз.

## § Описание

На долю патологии сетчатки приходится около 1% всех заболеваний глаза, хотя эта проблема является одной из самых актуальных в офтальмологии. Среди болезней сетчатки выделяют:

- сосудистые заболевания (диабетическая ретинопатия, непроходимость вен сетчатки, окклюзия центральной артерии сетчатки, изменение сетчатки при артериальной гипертензии);

- патологические изменения сетчатки (кровоизлияние, отек, макулярные разрывы сетчатки, новообразования сетчатки);

- травмы сетчатки (проникающие рубцы, инородные тела);

- расслоение сетчатки;

- дистрофические заболевания (возрастная макулодистрофия (дистрофия) сетчатки, центральная и периферическая дистрофия сетчатки при близорукости, пигментная дегенерация сетчатки);

- новообразования.

Наиболее распространенными патологиями (рисунок 1) сетчатки глаза по данным ВОЗ являются:

- хориоидальная неоваскуляризация (CNV);
- диабетический макулярный отек (DME);
- возрастная макулярная дегенерация (AMD);
- друзы (DRUSEN).

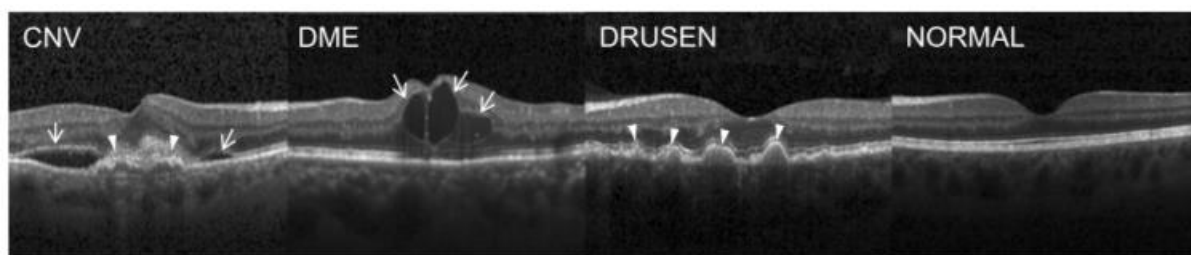


Рисунок 1 – Особенности томограмм для патологий (CNV, DME, Drusen) и здоровой сетчатки

## § Что было сделано

Стоит отметить, что для анализа медицинских изображений важно, чтобы при классификации, модель не отнесла пациента с патологией к группе здоровых.

Для оценки моделей описанных далее будут использоваться следующие метрики:

- точность;
- полнота;
- F1-мера.

Для обучения и оценки использовался следующий набор данных – «Labeled Optical Coherence Tomography (OCT) images for classification». Набор размеченных томограмм сетчатки содержит 83495 изображений для обучения и 968 изображений для оценки модели. Изображения разделены на 4 каталога: CNV, DME, DRUSEN и NORMAL [39]. Данные представлены изображениями разных размеров (512×496, 512×768, 1536×496) в формате .jpeg, которые затем были сжаты до размеров 227×227. Несмотря на сжатие почти в 2 раза, изображения обладают достаточной информативностью для классификации.

После сжатия набор был разделен в пропорции 80:20 на тренировочную и валидационную выборку, таким образом, получаем 66788 снимков для обучения и 16696 снимков для валидации. На данном наборе было обучено 7 моделей, описанных ниже, для всех моделей использовался оптимизатор “rmsprop” [40]. Для оценки потерь – категориальная перекрестная энтропия. Для обучения моделей использовалась библиотека «Tensorflow» и фреймворк «Keras».

## § Результаты и их сравнение

Первоначально была разработана CNN с 3-мя сверточными слоями, так как такого количества слоев достаточно для выделения признаков [20], кроме того изображения в наборе достаточно однородные, хотя и присутствуют выбросы в виде зашумленных изображений или изображений с белыми областями.

Из таблицы 1 видно, что относительно более глубоких моделей, линейная CNN с 3 сверточными слоями показывает на порядок более низкую точность. Затем, последовала разработка более глубоких моделей, например, таблица 1 показывает, что еще один свёрточный слой и увеличение количества фильтров позволяет кардинально улучшить долю правильных ответов. Блок-схемы с количеством слоёв, размерами фильтров, и другими параметрами моделей находятся в «.ipynb» файлах, вместе с матрицами неточностей, визуализированными картами активации, и визуализированные тепловые карты.

В таблице 1 видно, что модель с 5 свёрточными слоями дает еще более точные прогнозы в отношении диагноза, однако увеличение количества слоев с 5 до 6 не дает увеличения доли правильных ответов, что может говорить о том, что такая архитектура является слишком сложной для данной задачи, также это может сигнализировать о недообучении. Далее, к 6 линейным свёрточным слоям были добавлены дополнительные слои свертки с фильтром  $1 \times 1$ , таким образом получив модель с 10 сверточными слоями. Это дополнение дало увеличение точности на 1 процент, также из матрицы неточностей, видно, что данная модель выполнила основную задачу лучше остальных, так как ни один больной пациент не был классифицирован как здоровый. Далее были предприняты попытки разработать модель с Inception-модулями, так как Xception [40], является лидером соревнования «Imagenet», но использовать столь глубокую сеть для классификаций однородных томограмм было бы нецелесообразно с точки зрения затраченного времени на обучение. Следовательно, нужно было пойти на компромисс, либо уменьшить количество параметров – еще больше сжать изображения, либо разработать менее глубокую модель с такими модулями. Как можно видеть из таблицы 1 данная модель показывает себя хуже, чем линейная модель с 11 сверточными слоями. В ipython ноутбуках есть тепловые карты активации для каждого класса изображений, на которых можно увидеть, какие области изображений повлияли на принятие сетью решения, то есть с их помощью можно понять какие признаки важны для того или иного класса. При помощи

карт активаций, можно увидеть какие признаки выделялись в каждом слое. Для сравнения с более глубокими архитектурами была обучена модель «MobileNetV2», которая отличается сравнительно небольшим количеством параметров и высокой точностью [41]. Более детально изучив тепловые карты, можно сделать вывод, что модель с 3 слоями при классификации изображения класса CNV была «не уверена» в предсказании, так как на изображении отсутствуют желтые или красные области. Касательно других моделей – можно сделать вывод, что модели выделяли правильные области для классификации, например для класса DRUSEN в основном выделялась область с волнообразной деформации линии пигментного эпителия, также, для классификации здоровых сетчаток, алгоритмы «смотрят» на деформацию макулы.

При обучении использовались дополнительные функции “keras. callbacks” для предотвращения переобучения, поэтому количество эпох для обучения каждой модели разное. Все модели прекращали обучения если в течение трёх эпох обучения значение точности на валидации не изменялось более чем на  $\pm 0,1$ , далее, если значение функции потерь не изменялось в течение двух эпох, скорость обучения уменьшалась в 10 раз. Также из-за различий в сложности модели требуют разной вычислительной мощности, следовательно, нужно было корректировать параметр “batch size”, который равен количеству изображений передаваемых в нейросеть за раз, то есть, если “batch size” равен 32, то получим  $66788/32 = 2087$  шагов за одну эпоху.

## § Итоги

Таким образом были спроектированы алгоритмы анализа томограмм, использующие нейронные сети различной сложности: 3-слойные, 4-слойные, 5-слойные, 6-слойные, 10-слойные, а также модель с Inception модулями;

Также было показано, что наиболее точной является 10-слойная свёрточная нейронная сеть. Точность составляет 98.3% что сопоставимо с точностью «MobileNetV2» 98,7%, но разработанная модель требует значительно меньшего времени, а следовательно и ресурсов, на обучение.

Таблица 1 – Оценка моделей

Количество слоев	Доля правильных ответов, %	Класс	Точность	Полнота	F - мера
3	83,4	CNV	0,73	0,99	0,84
		DME	0,99	0,72	0,83
		Drusen	1,00	0,63	0,77
		Normal	0,78	1,00	0,88
4	94,5	CNV	0,92	0,98	0,95
		DME	0,98	0,86	0,92
		Drusen	1,00	0,94	0,97
		Normal	0,90	1,00	0,95
5	97,6	CNV	0,96	0,98	0,97
		DME	0,97	0,95	0,96
		Drusen	1,00	0,98	0,99
		Normal	0,97	1,00	0,98
6	97,0	CNV	0,92	1,00	0,95
		DME	1,00	0,95	0,97
		Drusen	1,00	0,93	0,96
		Normal	0,98	1,00	0,99
10	98,3	CNV	0,96	0,98	0,97
		DME	0,98	0,98	0,98
		Drusen	1,00	0,98	0,99
		Normal	1,00	1,00	1,00
Модель с модулями Inception	97,0	CNV	0,91	0,99	0,95
		DME	1,00	0,96	0,98
		Drusen	1,00	0,93	0,96
		Normal	0,99	1,00	0,99
MobileNetV2	98,7	CNV	0,99	1,00	0,97
		DME	1,00	1,00	0,98
		Drusen	1,00	0,99	0,99
		Normal	1,00	1,00	1,00