

Интеграция архитектуры на основе трансформеров и больших языковых моделей для анализа данных оптической когерентной томографии с целью повышения точности дифференциальной диагностики заболеваний сетчатки

Коньшина О.В. магистр, врач-офтальмолог ^{1,2}

Першин А.Д. аспирант, <https://orcid.org/0009-0007-7480-5054>¹

Кулябин М.К. к.т.н., <https://orcid.org/0009-0007-0440-030X>³

Борисов В.И. к.т.н., доцент <https://orcid.org/0000-0003-0486-7552>¹

1 ФГАОУ ВО «УрФУ имени первого Президента России Б.Н. Ельцина», ИРИТ–РТФ, ул. Мира 32, 620062, г. Екатеринбург

2 Клиника «МИК», ул. Кораблестроителей, д. 21. корп. 1, лит. В, 199226, г. Санкт-Петербург

3 ООО «ВизиоМедИИ», Головинское ш., д. 8, 125212, г. Москва

Ключевые слова: КЛАССИФИКАЦИЯ БИОМАРКЕРОВ ОКТ,

МУЛЬТИКЛАССОВАЯ КЛАССИФИКАЦИЯ ОКТ, LLM В

ОФТАЛЬМОЛОГИИ, DEEP LEARNING, ОСТ, RETINA, VIT, DEEPSEEK

Цель исследования: разработать и протестировать алгоритм взаимодействия между моделью глубокого обучения для мультиклассовой классификации биомаркеров на изображениях оптической когерентной томографии (ОКТ) и большой языковой моделью (LLM) DeepSeek-V3 с целью повышения точности дифференциальной диагностики заболеваний сетчатки.

Материалы и методы: нами были собраны и размечены два набора данных: тренировочный (n=3288 изображений ОКТ центральной зоны сетчатки, размеченных по 8 биомаркерам) и валидационный (n=50 клинических случаев с

сайта otcases.com). Для классификации биомаркеров сравнивались архитектуры ResNet, DenseNet, EfficientNet и Vision Transformer (ViT-Tiny-Patch16-224). Наибольшую эффективность показала модель ViT-Tiny с метриками F1 macro 0.84 ± 0.03 и ROC AUC 0.9. Разработан алгоритм интеграции полученных меток с данными клинического анамнеза через API-интерфейс LLM DeepSeek-V3.

Результаты: комбинация ОКТ-биомаркеров и анамнеза повысила диагностическую точность: Top-1 Accuracy – 78%, Top-3 – 94%, MRR – 84%, что на 10–44% выше, чем при использовании одного типа данных.

Заключение: предложенный подход позволяет автоматизировать выявление биомаркеров на ОКТ и повысить точность дифференциальной диагностики заболеваний глаз, обеспечивая снижение времени интерпретации снимков и поддержку клинических решений.

Введение.

Медицинская визуализация играет ключевую роль в современной офтальмологии, где особое значение имеет применение оптической когерентной томографии (ОКТ). ОКТ представляет собой неинвазивный метод визуализации, позволяющий проводить высокоточную оценку структур сетчатки на клеточном уровне с разрешением до 2–3 мкм [1]. Этот метод позволяет выявлять более 40 различных биомаркеров, характерных для широкого спектра заболеваний сетчатки [2].

В научной литературе широко представлены исследования, посвящённые классификации заболеваний сетчатки на основе ОКТ-изображений с использованием методов машинного и глубокого обучения [3–9]. Однако большинство этих исследований ограничены задачами четырёхклассовой классификации, тогда как в реальных клинических условиях число потенциальных диагнозов значительно превышает данное количество.

Альтернативный подход заключается в выявлении и классификации биомаркеров на изображениях ОКТ [10–13]. При этом один и тот же биомаркер может встречаться при различных заболеваниях, а один диагноз может проявляться различными наборами биомаркеров. Следовательно, для повышения точности

дифференциальной диагностики необходимо учитывать не только визуальные данные, но и клинический анамнез пациента.

В последние годы всё большее внимание уделяется применению больших языковых моделей (LLM), таких как GPT-4, Gemini, Grok и DeepSeek-V3, в медицинской практике [14–17]. Эти модели демонстрируют высокую эффективность в обработке естественного языка, генерации медицинских заключений и анализе сложных клинических сценариев. Тем не менее, их интеграция с данными визуализации требует дальнейших исследований.

Целью настоящего исследования является интеграция и тестирование алгоритма совместного использования нейросетевой модели для классификации биомаркеров на ОКТ и большой языковой модели DeepSeek-V3 с целью повышения точности дифференциальной диагностики заболеваний сетчатки.

Материалы и методы.

Наборы данных.

В нашем исследовании использовались два набора данных. Набор для обучения моделей для многоклассовой классификации был сформирован из трех источников: Открытые наборы OCTID [18] (946 изображений) и OCTDL [19] (2042 изображений), а также личный набор (300 изображений ОКТ макулярной зоны). Итоговый набор содержит 3288 изображений, размеченных по восьми

биомаркерам: IRF – интравитреальная жидкость, SRF – субретинальная жидкость, HE – твердые экссудаты, CNV – Субретинальные неоваскулярные мембраны всех типов + субретинальный гиперрефлективный материал, PED – элевация пигментного эпителия, Drusen – все типы друз, ERM - эпиретинальная мембрана, МН – объединение всех типов макулярных отверстий.

На рисунке 1 представлено распределение биомаркеров в тренировочном наборе.

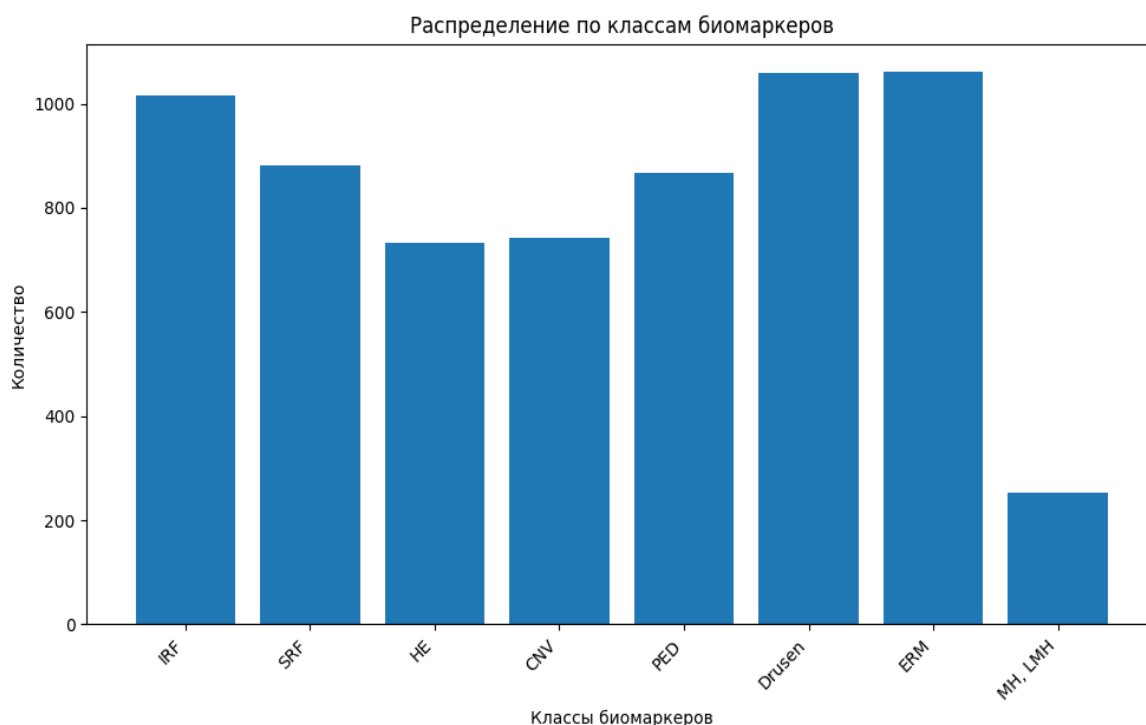


Рисунок – 1. Гистограмма распределения по классам биомаркеров.

В наборе данных сохраняется дисбаланс классов. Каждое изображение может содержать несколько биомаркеров.

В качестве тестового набора данных использованы аннотированные клинические случаи из открытого профессионального сайта OCT Cases (octcases.com), набор

содержит следующие данные: анамнез, изображение ОКТ и верифицированный диагноз. Всего в наборе представлено 18 разных диагнозов, часть из которых являются редко встречаемыми.

Предварительная обработка изображений осуществлялась с помощью библиотеки torchvision (PyTorch) и включала в себя следующий конвейер преобразований: изменение размера до 180 пикселей по меньшей стороне, не изменяя пропорции изображений при помощи функции Resize , случайные повороты (± 15 градусов) и горизонтальные отражения ($p=0.5$) (для увеличения разнообразия обучающих данных), перспективные преобразования ($p=0.3$, $distortion_scale=0.1$) (для имитации изменений в перспективе), преобразование в тензоры PyTorch, обрезка по центру до 224x224 пикселей (для соответствия размеру входа моделей) с добавлением автопадинга по меньшей стороне и нормализация ($mean=[0.485, 0.456, 0.406]$, $std=[0.229, 0.224, 0.225]$) (для ускорения обучения и улучшения сходимости. Эти значения соответствуют статистике ImageNet и используются для согласования с весами моделей). Этот конвейер обеспечил стандартизацию и аугментацию изображений, что способствовало повышению качества обучения моделей.

Обучение и выбор лучшей модели для мультиклассовой классификации биомаркеров на изображениях ОКТ сетчатки.

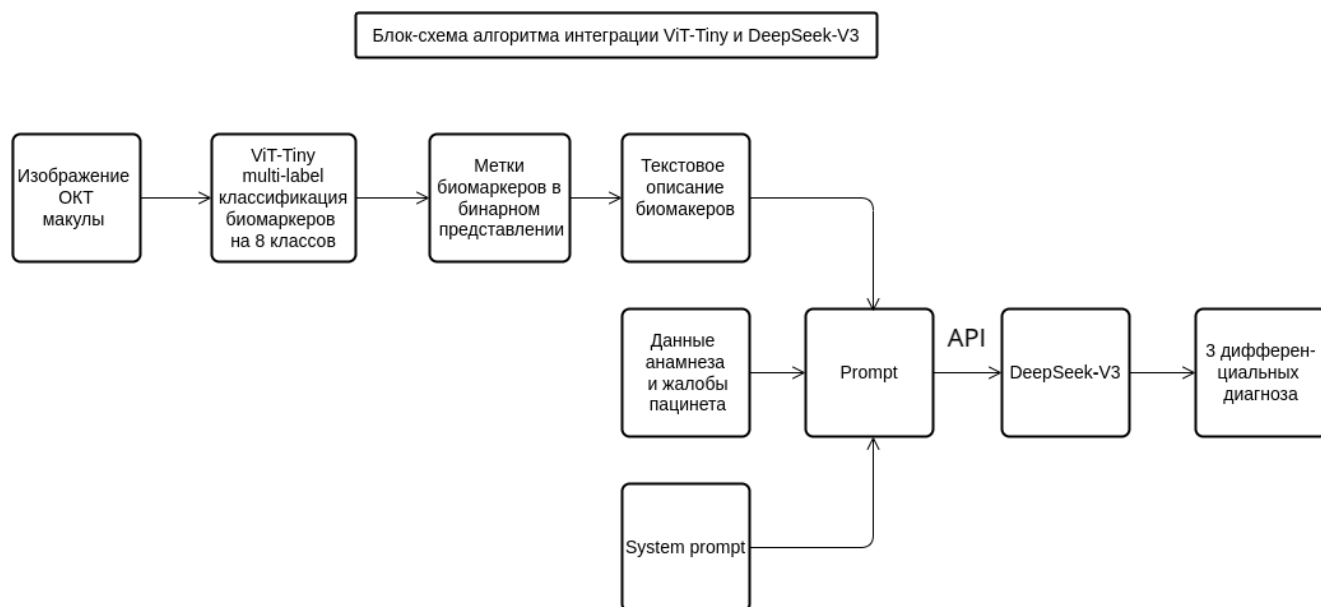
Для решения задачи multi-label классификации и изучения влияния различных архитектур на производительность, с учётом ограниченности данных, был сформирован набор моделей: ResNet50 и ResNet101 (сверточные нейронные сети с разной глубиной), DenseNet121 (с плотными связями между слоями), EfficientNet_b1 (с эффективным масштабированием) и ViT-Tiny-Patch16-224 (трансформерная архитектура). Этот выбор обеспечил разнообразие подходов и позволяет оценить их применимость к данной задаче при ограниченном объеме обучающих данных.

ViT-Tiny-Patch16-224 для классификации офтальмологических изображений с модификацией выходного слоя под 8 классов. Модель обучалась в течение 20 эпох с оптимизатором Adam и функцией потерь BCEWithLogitsLoss, используя пороговое значение 0.35 для бинаризации предсказаний. Для изменения скорости обучения Learning Rate использовалась схема “Warmup with Cosine Annealing”.

Для каждой модели вычислялись метрики Accuracy, Precision, Recall и F1-score (macro/micro averaging), с дополнительным тестированием на тестовом наборе данных. Для каждой метрики были построены доверительные интервалы 95% методом Бутстреп.

Валидация и взаимодействие с LLM.

Ниже представлена блок-схема алгоритма интеграции модели для multi-label классификации и LLM(DeepSeek-V3).



Рисунок– 2. Блок-схема алгоритма интеграции ViT-Tiny и DeepSeek-V3

Из полученных при помощи обученной модели ViT-Tiny-Patch16-224 меток биомаркеров и данных анамнеза был сформирован промт для взаимодействия с большой языковой моделью (LLM) DeepSeek-V3 через интерфейс программирования приложений (API). При этом метки на выходе из модели были бинарными для мультиклассовой классификации и были преобразованы в текстовое описание классов. Для оценки работы алгоритма было сформировано 3 варианта промта:

1. описание биомаркеров на снимках ОКТ и данные анамнеза;

2. только описание биомаркеров на снимках ОКТ;
3. только данные анамнеза.

Работа алгоритма была оценена при помощи следующих метрик Top-1 Accuracy, Top-3 Accuracy, Mean Reciprocal Rank (MRR).

Результаты.

В таблице 1 приведены полученные метрики качества обученных моделей для multi-label классификации биомаркеров на изображениях ОКТ сетчатки, рассчитанные на тестовом наборе данных.

Таблица – 1. метрики качества дообученных моделей для мультиклассовой классификации биомаркеров на изображениях ОКТ сетчатки.

Модель	F1 micro	F1 macro
ResNet50	0.81±0.02	0.81±0.02
ResNet 101	0.80 ± 0.03	0.80 ± 0.03
DenseNet121	0.75 ± 0.04	0.76 ± 0.04
EfficientNet_b1	0.64±0.04	0.61±0.04
ViT-Tiny	0.84 ± 0.02	0.84 ± 0.03

Лучший результат показала модель ViT-Tiny-Patch16-224. Метрика F1 Weighted для этой модели - 0.84 ± 0.04 , Specificity – 0.94 ± 0.01 .

Ниже представлена диаграмма сравнения метрик Precision, Recall и F1 для разных классов биомаркеров полученных на тестовых данных для модели ViT-Tiny-Patch16-224 (Рисунок 3).

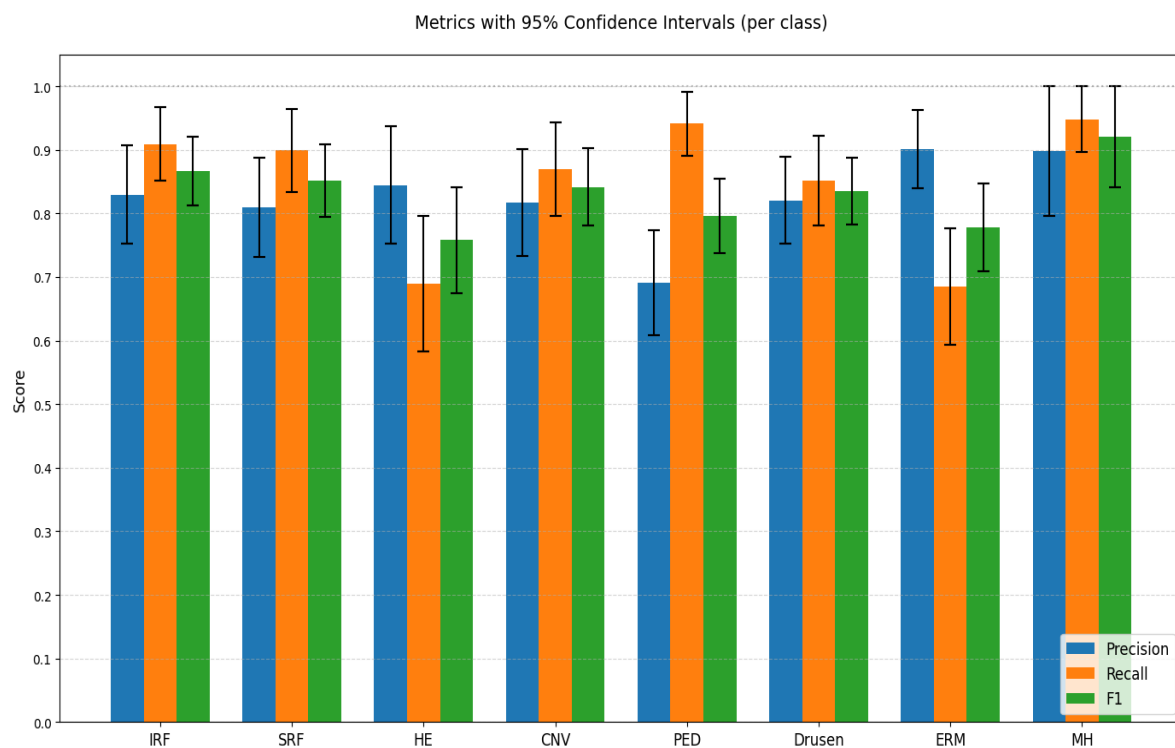


Рисунок – 3. Гистограмма сравнения метрик Precision, Recall и F1 для разных классов биомаркеров, полученных на тестовых данных для модели ViT-Tiny-Patch16-224.

На представленном графике видно, что наибольшие ошибки модели проявляются при классификации классов 2 (HE) и 6 (ERM), что, вероятно, обусловлено спецификой морфологии этих биомаркеров. Оба класса могут быть плохо

различимы на изображениях и подвержены потере информации в процессе предварительной обработки данных.

Далее представлен график ROC кривых для каждого класса биомаркеров, которая показывает способность модели различать классы. ROC-кривая строится для каждого класса отдельно и показывает, насколько хорошо модель разделяет положительные и отрицательные примеры. (Рисунок 4).

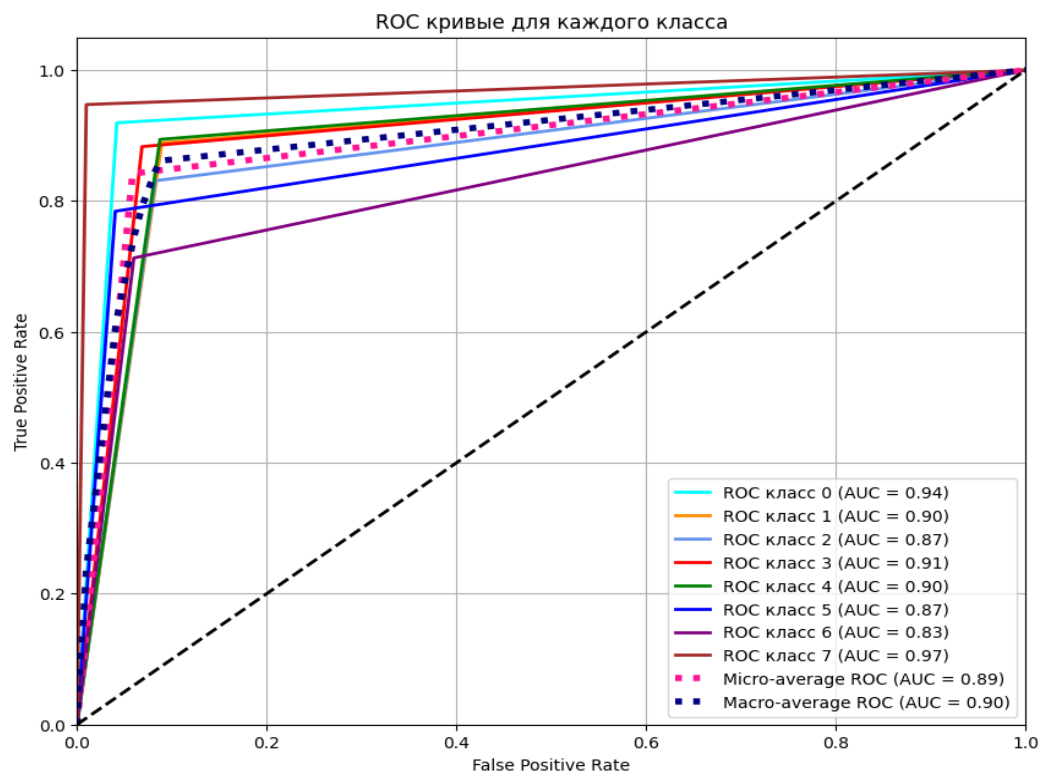


Рисунок 4. График ROC кривых для каждого класса биомаркеров.

На представленном графике понятно, что все классы, кроме ERM, имеют $AUC > 0.87$ (хорошая сбалансированность). Macro-average 0.90 подтверждает надежность модели для мультиклассовой диагностики.

Интерпретация работы модели ViT-Tiny-Patch16-224 с использованием Grad-CAM++.

Grad-CAM (Grad-weighted Class Activation Mapping) – это метод интерпретации моделей глубокого обучения, который позволяет визуализировать важные области изображения, используемые моделью для принятия решений. Grad-CAM++ – это улучшенная версия этого метода, предлагающая более точные и детализированные карты активации [20].

На рисунке 5 приведены примеры карт активации для разных классов биомаркеров.

Модель фокусируется на определенных областях изображения, что может быть полезно для понимания того, как модель принимает решения.

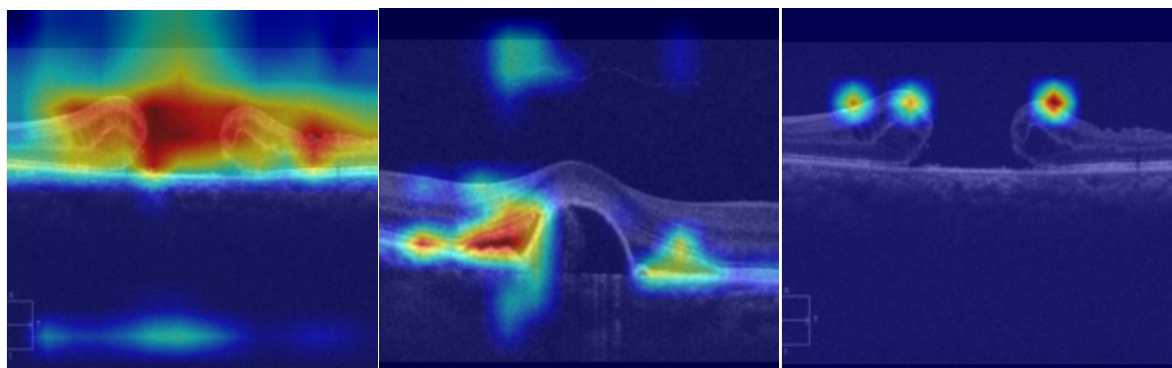


Рисунок 5. Карты активации для разных классов биомаркеров. (Слева на право классы “MH”, “SRF”, “ERM”)

Оценка алгоритма взаимодействия дообученной модели ViT-Tiny-Patch16-224 и LLM (DeepSeek-V3) в дифференциальной диагностике заболеваний сетчатки представлена в таблице 2.

Таблица 2. Метрики качества работы LLM (DeepSeek-V3) с разными вариантами промта.

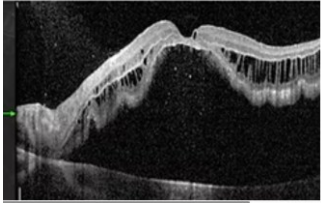
метрика/ промт	ОКТ+анамнез	ОКТ	Анамнез
Top-1 Accuracy	0.78 ±0.11	0.52± 0.14	0.30± 0.14
Top-3 Accuracy	0.94±0.8	0.66±0.13	0.46±0.14
MRR	0.84±0.04	0.58±0.06	0.41±0.06

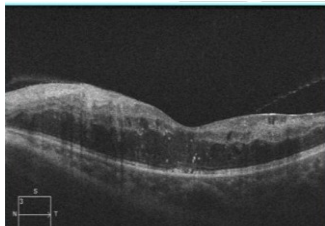
Исходя из полученных данных можно сделать следующие вывод: наилучшие результаты достигаются при использовании комбинации данных, полученных с помощью нейронной сети Vit-Tiny (дообученной для задачи мультиклассовой классификации биомаркеров на изображениях ОКТ) и текстовых данных анамнеза.

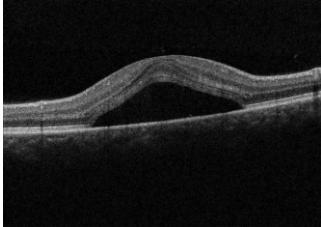
Ниже представлены примеры промта и работы алгоритма в таблице 3.

Таблица–3. Примеры работы алгоритма с разными вариантами промта.

№	изображение ОКТ макулы.	Выделенные биомаркеры	Варианты промта	результат работы алгоритма (3 дифференциальных диагноза)	верифицированный диагноз.

1.		IRF, SRF	ОКТ	Ты врач-офтальмолог клинический пример: на снимке ОСТ макулярной зоны выявлено: IRF, SRF. Напиши кратко 3 наиболее вероятных диагноза. Только названия клинических диагнозов.	1. Возрастная макулярная дегенерация (ВМД). 2. Центральная серозная хориоретинопатия (ЦСХРП). 3. Диабетический макулярный отек (ДМО).	Отслойка сетчатки
			Анамнез	Ты врач-офтальмолог клинический пример: 11-летний мальчик-афак с историей микросферофакии s/p операции по удалению катаракты поступил в отделение неотложной помощи с жалобами на снижение зрения на левый глаз в течение 3 месяцев. BCVA была 20/40 OD и 20/100 OS. Его зрачки оставались мейотическими, несмотря на многочисленные попытки расширить их фармакологическими средствами. Напиши кратко 3 наиболее вероятных диагноза. Только названия клинических диагнозов.	1. Вторичная катаракта (фиброз задней капсулы). 2. Зрачковый блок (глаукома) 3. Синдром микрофакии с вторичными осложнениями.	
			ОКТ + Анамнез	Ты врач-офтальмолог клинический пример: 11-летний мальчик-афак с историей микросферофакии s/p операции по удалению катаракты поступил в отделение неотложной помощи с жалобами на снижение зрения на левый глаз в течение 3 месяцев. BCVA была 20/40 OD и 20/100 OS. Его зрачки оставались мейотическими, несмотря на многочисленные попытки расширить их фармакологическими средствами. на снимке ОСТ макулярной зоны выявлено: IRF, SRF. Напиши кратко 3 наиболее вероятных диагноза. Только названия	1. Макулярный отек 2. Отслойка сетчатки (серозная или экссудативная) 3. Хронический увеит с осложнениями	

				клинических диагнозов.		
2.		IRF, HE	ОКТ	Ты врач-офтальмолог клинический пример: на снимке ОСТ макулярной зоны выявлено: IRF, HE. Напиши кратко 3 наиболее вероятных диагноза. Только названия клинических диагнозов.	1. Возрастная макулярная дегенерация (ВМД). 2. Центральная серозная хориоретинопатия (ЦСХРП). 3. Диабетический макулярный отек (ДМО).	Радиационная ретинопатия.
			Анамнез	Ты врач-офтальмолог клинический пример: 24-летний мужчина был направлен из-за безболезненного нечеткого зрения (OD > OS) в течение 1 месяца. Его прошлый медицинский анамнез был значительным в отношении опухоли пинеальной железы, которая лечилась краниоспинальной лучевой терапией (36 Гр, 54 Гр, а затем 59,4 Гр). У него не было диабета или гипертонии. При осмотре острота зрения составляла 20/150 OD и 20/40-1 OS. Цветовое зрение составляло 0/12 OD и 7/14 OS. Напиши кратко 3 наиболее вероятных диагноза. Только названия клинических диагнозов.	1. Радиационная ретинопатия. 2. Радиационная оптическая нейропатия. 3. Рецидив опухоли пинеальной области с компрессией зрительных путей.	
			ОКТ + Анамнез	Ты врач-офтальмолог клинический пример: 24-летний мужчина был направлен из-за безболезненного нечеткого зрения (OD > OS) в течение 1 месяца. Его прошлый медицинский анамнез был значительным в отношении опухоли пинеальной железы, которая лечилась краниоспинальной лучевой терапией (36 Гр, 54 Гр, а затем 59,4 Гр). У него не было диабета или гипертонии. При осмотре	1. Радиационная ретинопатия. 2. Хориоретинопатия (вторичная на фоне лучевой терапии). 3. Оклюзия вены сетчатки (возможно, лучевого генеза).	

				острота зрения составляла 20/150 OD и 20/40-1 OS. Цветовое зрение составляло 0/12 OD и 7/14 OS. на снимке OCT макулярной зоны выявлено: IRF, HE. Напиши кратко 3 наиболее вероятных диагноза. Только названия клинических диагнозов.		
3.		SRF	ОКТ	Ты врач-офтальмолог клинический пример: на снимке OCT макулярной зоны выявлено: SRF. Напиши кратко 3 наиболее вероятных диагноза. Только названия клинических диагнозов.	1. Центральная серозная хориоретинопатия (ЦСХ) 2. Возрастная макулярная дегенерация (ВМД) 3. Хориоидальная неоваскуляризация (ХНВ).	Центральная серозная хориоретинопатия (ЦСХ).
			Анам нез	Ты врач-офтальмолог клинический пример: 27-летний мужчина, спортсмен, обратился в связи с нечетким зрением на правом глазу, связывает со стрессом. Напиши кратко 3 наиболее вероятных диагноза. Только названия клинических диагнозов.	1. Миопия (близорукость). 2. Кератоконус. 3. Отслойка сетчатки.	
			ОКТ + Анам нез	Ты врач-офтальмолог клинический пример: 27-летний мужчина, спортсмен, обратился в связи с нечетким зрением на правом глазу, связывает со стрессом. на снимке OCT макулярной зоны выявлено: SRF. Напиши кратко 3 наиболее вероятных диагноза. Только названия клинических диагнозов.	1. Центральная серозная хориоретинопатия (ЦСХ). 2. Возрастная макулярная дегенерация (ВМД) (редко в 27 лет, но возможно при ранних формах). 3. Хориоидальная неоваскуляризация (ХНВ) (идиопатическая или на фоне миопии).	

Обсуждение.

В рамках проведения данного исследования была разработана и протестирована модель мультиклассовой классификации восьми биомаркеров на изображениях ОКТ сетчатки на основе предобученной модели ViT-Tiny-Patch16-224. Именно эта модель демонстрирует наилучшие метрики качества работы, сопоставимые с данными научной литературы.

В таблице 3 приведено сравнение результатов качества алгоритмов глубокого обучения.

Таблица – 4. Сравнение результатов качества алгоритмов глубокого обучения.

№	Статья	Алгоритм	Метрики качества.
1.	Настоящее исследование (кол-во классов $n = 8$)	ViT-Tiny-Patch16-224	Recall – 83,38% (0.83 ± 0.02) Precision – 86.2% (0.86 ± 0.05) Specificity – 94% (0.94 ± 0.01)
2.	Ali Salehi и др[11] (кол-во классов $n = 8$)	ResNet50	Recall – 86.8%, Specificity – 87.4%
3.	Каталевская Е.А. и др[10] (кол-во классов $n = 6$)	EfficientNetB0 + FPN (Feature Pyramid Network), задача сегментации	Recall от 82,64% до 94,44%, Specificity от 89,73% до 99,41%

Таким образом, использование известной архитектуры ViT-Tiny для решения задачи мультиклассовой классификации позволило достичь качества работы алгоритма, сравнимого с результатами ранее опубликованных исследований.

В таблице 4 представлены результаты трех исследований, касающихся применения различных LLM в медицинской сфере (офтальмология и дерматология). В каждом случае используется разный подход и метрики, что затрудняет прямое сравнение, но позволяет оценить относительную эффективность каждого метода и предметной области. Размер выборок (n) варьируется, что влияет на статистическую значимость результатов.

Таблица – 5. Сравнения различных LLM в разных областях медицины.

№	Статья	Алгоритм	Метрики качества.
1.	Настоящее исследование(n=50), офтальмология	Vit-Tiny + DeepSeek	Top 1 – 78% Top 3 – 94% MRR – 84%
2.	Цзюэсяо Чжоу и др. (n=150), [17] дерматология	SkinGPT-4(Vit+Gpt 4.0)	80,63% – были оценены сертифицированными дерматологами как правильные или релевантные. Эта оценка охватывала как полностью согласен – 75,00%
3.	D. Bahir, и др[14] (n=600), офтальмология	ChatGPT-4, Gemini	GeminiAdvanced – точность 66%. ChatGPT-4 – точность 62%

Несмотря на методологические различия между исследованиями, позволяющими лишь качественное сравнение полученных результатов, можно утверждать, что

модели на основе больших языковых моделей (LLM), такие как SkinGPT-4, ChatGPT-4, Gemini и DeepSeek-V3, демонстрируют высокую эффективность в решении задач медицинской диагностики. Особенно перспективным представляется интеграция Vision Transformer и LLM, как это реализовано в архитектурах ViT-Tiny + DeepSeek и SkinGPT-4. Для более строгой оценки и объективного сравнения различных подходов необходимы дальнейшие исследования, основанные на унифицированных метриках качества, стандартизированных наборах данных и воспроизводимых экспериментальных протоколах.

Заключение.

В ходе выполнения настоящего исследования была оценено качество работы 5 нейронных сетей различной архитектуры для решения задачи мультиклассовой классификации и получены следующие значения: ResNet 50 F1 macro – 0.81 ± 0.02 , ResNet101 – F1 macro – 0.80 ± 0.03 , DenseNet – F1 macro 0.76 ± 0.04 , Efficient Net B1 – F1 macro 0.61 ± 0.04 , ViT-Tiny-Patch 16-224 – F1 macro 0.84 ± 0.03 . Лучшей моделью оказалась архитектура ViT-Tiny-Patch16-224. Для данной модели были рассчитаны дополнительные метрики качества ROC AUC macro – 0.9, Recall (чувствительность) средняя – 83,38%, Precision средняя – 86.2 %, специфичность – 94%.

Следующим этапом разработан алгоритм взаимодействия LLM (DeerSeek-V3) по API. Было использовано 3 варианта промта сформированных из валидационного набора данных и меток биомаркеров полученных при помощи обученной модели ViT-Tiny-Patch16-224 (ОКТ биомаркеры и текстовые данные анамнеза, только текстовые данные анамнеза, только метки ОКТ биомаркеров). В качестве предсказания DeerSeek были получены три дифференциальных диагноза. Эти диагнозы сопоставлялись с верифицированными диагнозами из набора. Лучшее качество предсказаний получилось при использовании полных данных (метки ОКТ и текстовые данные анамнеза). Для оценки работы алгоритма использовались следующие метрики: Top-1 – 78%, Top-3 – 94%, MRR – 84% – что показывает эффективность работы алгоритма. При помощи данного алгоритма удалось улучшить метрики Top-1, Top-3 и MRR более чем на 10%, по сравнению с методами, использующими только клинический анамнез, или данные о наличии биомаркеров на снимках ОКТ сетчатки (на 50%, 42% и 44% соответственно).

Финансирование исследования. Исследование выполнено при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках Программы развития Уральского федерального университета имени первого Президента России Б.Н. Ельцина в соответствии с программой стратегического академического лидерства "Приоритет-2030".

Конфликт интересов отсутствует.

Литература/References.

1. Использование оптической когерентной томографии в диагностике заболеваний сетчатки (обзор литературы) [Интернет]. [цитируется по 3 март 2025 г.]. Доступно на: <https://cyberleninka.ru/article/n/ispolzovanie-opticheskoy-kogerentnoy-tomografii-v-diagnostike-zabolevaniy-setchatki-obzor-literatury/viewera>.
2. Pandya BU, Grinton M, Mandelcorn ED, Felfeli T. RETINAL OPTICAL COHERENCE TOMOGRAPHY IMAGING BIOMARKERS: A Review of the Literature. RETINA. 2024 Mar;44(3):369.
3. Yang M, Du J, Lv R. CRAT: Advanced transformer-based deep learning algorithms in OCT image classification. Biomedical Signal Processing and Control. 2025 Jun 1;104:107544.
4. Wang X, Tang F, Chen H, Luo L, Tang Z, Ran AR, et al. UD-MIL: Uncertainty-Driven Deep Multiple Instance Learning for OCT Image Classification. IEEE J Biomed Health Inform. 2020 Dec;24(12):3431–42.
5. Üzen H, Fırat H, Alperen Özçelik ST, Yusufoglu E, Çiçek İB, Şengür A. Central serous retinopathy classification with deep learning-based multilevel feature extraction from optical coherence tomography images. Optics & Laser Technology. 2025 Jun 1;184:112519.
6. Khudhur AM. Optical coherence tomography retinal classification using deep neural network. J Opt [Internet]. 2025 Feb 1 [cited 2025 Jul 9]; Available from: <https://doi.org/10.1007/s12596-025-02561-6>.
7. Abbas Y, Hadi HJ, Aziz K, Ahmed N, Akhtar MU, Alshara MA, et al. Reinforcement-based leveraging transfer learning for multiclass optical coherence tomography images classification. Sci Rep. 2025 Feb 20;15(1):6193.
8. Diao S, Su J, Yang C, Zhu W, Xiang D, Chen X, et al. Classification and segmentation of OCT images for age-related macular degeneration based on dual guidance networks. Biomedical Signal Processing and Control. 2023 Jul 1;84:104810.

9. Hamid L, Elnokrashy A, Abdelhay EH, Abdelsalam MM. A deep learning LSTM-based approach for AMD classification using OCT images. *Neural Comput & Applic.* 2024 Nov 1;36(31):19531–47.
10. Каталевская Е.А., Сизов А.Ю., Гилемзянова Л.И. Алгоритм искусственного интеллекта для сегментации патологических структур на сканах оптической когерентной томографии сетчатки глаза | Российский журнал телемедицины и электронного здравоохранения [Internet]. [cited 2025 Mar 3]. Available from: <https://jtelemed.ru/article/algorithm-iskusstvennogo-intellekta-dlja-segmentacii-patologicheskikh-struktur-na-skanah-optic>.
11. Salehi A, Lee GC, Ren H, Yu S, Singh R, Talcott K, et al. Multi-pathology classification of retinal OCT scans. *Investigative Ophthalmology & Visual Science.* 2023 Jun 1;64(8):1086.
12. Arzamastsev AA, Анатольевич AA, Arzamastsev AA, Fabrikantov OL, Львович ФО, Fabrikantov OL, et al. Classification of optical coherence tomography images using deep machine-learning methods. *Digital Diagnostics.* 2024 Apr 19;5(1):5–16.
13. Kulyabin M, Zhdanov A, Pershin A, Sokolov G, Nikiforova A, Ronkin M, et al. Segment Anything in Optical Coherence Tomography: SAM 2 for Volumetric Segmentation of Retinal Biomarkers. *Bioengineering.* 2024 Sep;11(9):940.
14. Bahir D, Zur O, Attal L, Nujeidat Z, Knaanie A, Pikkell J, et al. Gemini AI vs. ChatGPT: A comprehensive examination alongside ophthalmology residents in medical knowledge. *Graefes Arch Clin Exp Ophthalmol.* 2025 Feb 1;263(2):527–36.
15. ResearchGate [Internet]. 2024 [cited 2025 Mar 2]. (PDF) BiMediX2: Bio-Medical EXpert LMM for Diverse Medical Modalities. Available from: https://www.researchgate.net/publication/386949414_BiMediX2_Bio-Medical_EXpert_LMM_for_Diverse_Medical_Modalities.
16. Li C, Wong C, Zhang S, Usuyama N, Liu H, Yang J, et al. LLaVA-Med: Training a Large Language-and-Vision Assistant for Biomedicine in One Day.

17. Zhou J, He X, Sun L, Xu J, Chen X, Chu Y, et al. SkinGPT-4: An Interactive Dermatology Diagnostic System with Visual Large Language Model [Internet]. medRxiv; 2023 [cited 2025 Mar 4]. p. 2023.06.10.23291127. Available from: <https://www.medrxiv.org/content/10.1101/2023.06.10.23291127v1>
18. Gholami P, Roy P, Parthasarathy MK, Lakshminarayanan V. OCTID: Optical coherence tomography image database. Computers & Electrical Engineering. 2020 Jan 1;81:106532.
19. Kulyabin M, Zhdanov A, Nikiforova A, Stepichev A, Kuznetsova A, Ronkin M, et al. OCTDL: Optical Coherence Tomography Dataset for Image-Based Deep Learning Methods. Sci Data. 2024 Apr 11;11(1):365.
20. Yamauchi T. Spatial Sensitive Grad-CAM++: Improved Visual Explanation for Object Detectors via Weighted Combination of Gradient Map. In 2024 [cited 2025 Jul 9]. p. 8164–8. Available from: https://openaccess.thecvf.com/content/CVPR2024W/XAI4CV/html/Yamauchi_Spatial_Sensitive_Grad-CAM_Improved_Visual_Explanation_for_Object_Detectors_via_CVPRW_2024_paper.html.