



Юсеф Ю.Н.¹, Бурсов А.И.^{2,3} ✉, Плюхова А.А.¹, Барышев К.В.¹, Смолин С.А.¹,
Сидамонидзе А.¹, Старцев Н.С.⁴

¹ Научно-исследовательский институт глазных болезней имени М.М. Краснова,
Москва, Россия

² Институт системного программирования имени В.П. Иванникова Российской
академии наук, Москва, Россия

³ Российский университет дружбы народов имени Патриса Лумумбы, Москва,
Россия

⁴ ООО «Круг», Москва, Россия

Модель глубокого обучения с улучшенной интерпретацией для детекции патологических признаков возрастной макулярной дегенерации на ОКТ

Конфликт интересов: не заявлен.

Вклад авторов: Юсеф Ю.Н. – общее руководство исследованием, формирование научной концепции и контроль выполнения работы; Бурсов А.И. – координация исследования, организационное сопровождение и взаимодействие между участниками проекта, подготовка текста рукописи, литературный обзор; Плюхова А.А. – координация исследования, организационное сопровождение и взаимодействие между участниками проекта, научная редакция; Барышев К.В. – сбор данных, их первичная обработка и аннотирование; Смолин С.А. – участие в сборе данных и их аннотировании; Сидамонидзе А. – разметка и контроль качества аннотаций; Старцев Н.С. – разработка, обучение и валидация моделей искусственного интеллекта.

Подана: 13.02.2026

Принята: 11.05.2026

Контакты: bursov@ispras.ru

Резюме

Введение. Автоматизированный анализ изображений оптической когерентной томографии (ОКТ) играет ключевую роль в ранней диагностике и мониторинге возрастной макулярной дегенерации (ВМД). Использование алгоритмов машинного обучения (искусственный интеллект – ИИ) открывает большие перспективы в ранней диагностике заболеваний сетчатки.

Цель. Создание интерпретируемой модели глубокого обучения для детекции патологических признаков ВМД на срезах ОКТ.

Материалы и методы. В отличие от традиционных методов, использующих жесткие бинарные метки, были применены мягкие метки (признаки ВМД), отражающие степень согласия врачей-экспертов при аннотировании. Это позволило учесть диагностическую неопределенность и снизить влияние межэкспертного разброса при аннотировании. Обучение проводилось на моделях семейства ResNet и ConvNeXtV2 Base и Large, предварительно предобученных на ImageNet (масштабная база данных аннотированных изображений, предназначенная для использования в исследованиях по распознаванию визуальных объектов. Содержит более 14 миллионов изображений высокого разрешения, охватывающих тысячи категорий объектов).

Результаты. Результаты показали высокую эффективность модели: средние значения чувствительности – 0,877, специфичности – 0,966, F1-меры – 0,880. Особенно



высокие показатели достигнуты для таких признаков, как интравитреальные кисты и отслойки ($F1 > 0,90$). Критически важным оказалось улучшение интерпретируемости: визуализация с помощью Grad-CAM продемонстрировала, что модель фокусируется на анатомически релевантных зонах патологии, а не на фоновых артефактах.

Заключение. Предложенный подход повышает как точность, так и доверие к решениям модели, что делает ее перспективным инструментом для вспомогательной диагностики в клинической офтальмологии.

Ключевые слова: оптическая когерентная томография, возрастная макулярная дегенерация, глубокое обучение, нейронные сети, ConvNeXt, soft labels, интерпретируемость моделей, Grad-CAM, автоматическая диагностика, анализ медицинских изображений

Yusef Y.¹, Bursov A.^{2,3} ✉, Plyukhova A.¹, Baryshev K.¹, Smolin S.¹, Sidamonidze A.¹, Startsev N.⁴

¹ Krasnov Research Institute of Eye Diseases, Moscow, Russia

² Ivannikov Institute for System Programming of the Russian Academy of Sciences, Moscow, Russia

³ Peoples' Friendship University of Russia named after Patrice Lumumba, Moscow, Russia

⁴ "Krug" LCC, Moscow, Russia

Deep Learning Model with Improved Interpretation for Detecting Pathological Signs of Age-Related Macular Degeneration on OCT

Conflict of interest: nothing to declare.

Authors' contribution: Yusef Yu. – general management of the study, formation of the scientific concept and supervision of the work; Bursov A. – study coordination, organizational support and interaction between the project participants, preparation of the manuscript, literature review; Plyukhova A. – study coordination, organizational support and interaction between the project participants, scientific editing; Baryshev K. – data collection, their primary processing and annotation; Smolin S. – participation in data collection and their annotation; Sidamonidze A. – markup and quality control of annotations; Startsev N. – development, training and validation of artificial intelligence models.

Submitted: 13.02.2026

Accepted: 11.05.2026

Contacts: bursov@ispras.ru

Abstract

Introduction. Automated optical coherence tomography (OCT) image analysis plays a key role in the early diagnosis and monitoring of age-related macular degeneration (AMD). The use of machine learning algorithms (Artificial intelligence, AI) opens up great prospects in the early diagnosis of retinal diseases.

Purpose. Creation of an interpreted deep learning model for detecting pathological signs of AMD in OCT slices.

Materials and methods. Unlike traditional methods using hard binary labels, soft labels (signs of AMD) were used, reflecting the degree of agreement of expert doctors during annotation. This made it possible to take into account diagnostic uncertainty and reduce the impact of cross-expert variation during annotation. The training was conducted on models from the ResNet and ConvNeXtV2 Base and Large families, pre-trained on

ImageNet (a large-scale database of annotated images intended for use in research on visual object recognition. Contains more than 14 million high-resolution images covering thousands of object categories).

Results. The results showed high efficiency of the model: average values of sensitivity – 0.877, specificity – 0.966, F1-measures – 0.880. Especially high values were achieved for such signs as intraretinal cysts and detachment ($F1 > 0.90$). Improvement of interpretability was critically important: visualization using Grad-CAM demonstrated that the model focuses on anatomically relevant areas of pathology, rather than on background artifacts.

Conclusion. The proposed approach increases both the accuracy and confidence in the model's solutions, which makes it a promising tool for auxiliary diagnostics in clinical ophthalmology.

Keywords: optical coherence tomography, age-related macular degeneration, deep learning, neural networks, ConvNeXt, soft labels, interpretability of models, Grad-CAM, automatic diagnostics, medical image analysis

■ ВВЕДЕНИЕ

Оптическая когерентная томография (ОКТ) произвела революцию в офтальмологии, являясь в настоящее время одним из ключевых методов диагностики и позволяя с высоким разрешением визуализировать слои сетчатки.

ОКТ с большой степенью достоверности помогает в диагностике и мониторинге заболеваний глазного дна, в том числе возрастной макулярной дегенерации, своевременно выявляя мельчайшие изменения – биомаркеры в слоях сетчатки, которые имеют критическое значение в прогнозе заболевания. Этот неинвазивный метод визуализации позволяет проводить раннюю диагностику и разрабатывать индивидуальные стратегии лечения, что в итоге способствует сохранению зрительных функций и, соответственно, трудоспособности и социальной адаптации.

Однако непосредственно анализ данных ОКТ сопряжен с рядом проблем, включая управление полученными снимками, влияние параметров и интерпретацию изображений. Эти проблемы требуют инновационных решений для повышения точности диагностики и клинической эффективности.

Данные ОСТ из разных источников могут быть противоречивыми, что усложняет анализ. Унифицированный интерфейс, структурирующий и интегрирующий данные с медицинскими картами, дает возможность упростить исследования и улучшить результаты диагностики.

Инструменты визуальной аналитики, позволяющие изучать и подчеркивать конкретные аспекты, помогают снизить сложность данных и сосредоточить внимание на соответствующих подмножествах, что способствует более точным оценкам [1].

Нечеткость изображений, часто встречающаяся на снимках ОСТ, может привести к ошибкам в измерениях. Исправление пространственных соотношений сторон путем использования нерастянутых изображений или применения математических поправок имеет решающее значение для повышения точности работы модели [2].

Глубокое обучение предлагает перспективные решения для автоматизированного анализа изображений ОКТ, включая сегментацию, классификацию и прогнозирование патологий. Однако остаются нерешенными следующие задачи: дефицит



данных, прозрачность моделей и различия в производительности в реальных условиях [3].

Методологические сложности глубокого обучения, такие как нестандартизированные протоколы визуализации и ограниченные вычислительные ресурсы, препятствуют обобщению моделей искусственного интеллекта на разных устройствах, в разных группах населения. С целью повышения эффективности применения ИИ в клинической практике необходимо решение данных проблем [4].

Хотя аппараты ОКТ продолжают совершенствоваться, проблемы с анализом полученных результатов сохраняются. Решение этих задач за счет улучшения управления данными, анализа параметров и использования искусственного интеллекта может расширить диагностические возможности ОСТ. Однако интеграция решений в клиническую практику требует тщательного рассмотрения нормативных и технических стандартов для обеспечения широкого распространения и повышения эффективности [5].

■ МАТЕРИАЛЫ И МЕТОДЫ

В качестве источника информации использовались данные, полученные в ФГБНУ «НИИГБ им. М.М. Краснова» с 3 приборов производства компании HEIDELBERG (Spectralis HRA+ОСТ (Heidelberg), Германия). Ввиду ограничений производителя экспорт данных с приборов производился вручную, в дальнейшем они собирались верхнеуровнево в формате E2E и делились на 2 категории: «неоваскулярная форма ВМД» и «ВМД сухая форма». На всех приборах использовалась одна программа для получения снимков ОКТ – Neurex.

Далее каждое исследование анонимизировалось и конвертировалось в серии изображений в формате PNG с помощью открытого пакета Python `oct_converter`, использованного в самописном скрипте. Каждая серия состояла из 24 и более файлов. Отличительной особенностью ВМД является то, что не все снимки в серии действительно содержат патологические признаки. Поэтому для аннотирования изображений (в данном контексте аннотирование – это трансляция врачебного опыта в машиночитаемый формат) была написана специализированная система в виде веб-приложения, с помощью которой врачи-эксперты с опытом от 3 лет работы и более по специальности могли отмечать наличие или отсутствие признаков ВМД на каждом снимке. Интерфейс системы представлен на рисунке.

Важным аспектом подхода к аннотированию служило то, что для каждого изображения было получено по 3 аннотации от группы врачей, т. е. для повышения достоверности 1 изображение показывалось 3 независимым экспертам. Это влияет на скорость процесса аннотирования, но повышает его качество.

Были выбраны следующие признаки для разметки изображений: «Атрофия ретинального пигментного эпителия», «Друзы», «Изменения по типу отека», «Инtrarетинальные кисты», «Отслойка нейроэпителия», «Отслойка ретинального пигментного эпителия». Если на изображении не было выявлено признаков ВМД, врачи выбирали «Признаки не найдены». При возникновении проблем во время аннотирования эксперты выбирали один или оба из следующих признаков: «Нужны дополнительные снимки для заключения», «Снимок непригоден для анализа».

Признаки отбирались по принципу наибольшей визуальной идентификации и патогномичности. Каждый из признаков указывает на стадию заболевания, по



Интерфейс приложения для аннотирования изображений с признаками ВМД
Application interface for annotating images with AMD features

которой можно определить дальнейшую тактику ведения пациента. При сочетании ряда признаков может быть поставлен диагноз неоваскулярной формы ВМД и своевременно рекомендована анти-VEGF-терапия.

После завершения разметки данных изображения и результаты аннотирования в анонимизированном формате выгружались в архив и передавались для дальнейшего анализа и обучения моделей искусственного интеллекта (ИИ).

Всего в выборке было более 42 200 снимков с признаками ВМД, из них 53% относилось к влажной форме и 47% – к сухой.

Дополнительно использовалась база из 6612 снимков без разделения по признакам, но с бинарным разделением – есть/нет признаков соответствующей формы на изображении. Среди этих изображений 1353 были с признаками сухой формы, 914 – без признаков сухой формы, 2204 – с признаками влажной формы, 2140 – без признаков влажной формы. Эта база использовалась для первичной валидации моделей.

Исследователи использовали 2 современные нейронные сети – ConvNeXt Base и ConvNeXt Large, которые изначально были обучены распознавать изображения на большой базе данных ImageNet. Такой подход является общепринятым в различных сферах, где применяется машинное обучение, и он значительно экономит время и ресурсы, необходимые для обучения моделей.

ConvNeXt Base и ConvNeXt Large применялись для анализа изображений с целью определения наличия определенных признаков (наличие/отсутствие сухой/влажной формы ВМД). Каждый признак оценивался просто как «есть» (1) или «нет» (0), без учета того, насколько уверенно с этим согласились эксперты-врачи.

В результате полученная модель положительно справлялась с задачей – по метрике F1 (это способ оценить точность работы модели, особенно когда важно учитывать и правильные находки, и ошибки, объединяет 2 характеристики: точность и полноту) она давала приемлемые результаты для отдельных признаков. Однако с помощью метода визуализации Grad-CAM стало понятно, на какие именно участки



изображения модель обращает внимание при принятии решения, оказалось, что она часто фокусируется на нерелевантных, не имеющих отношения к локализации патологии областях. Это значит, что, даже если ответ был правильным, модель могла прийти к нему неправильным путем, что ставит под сомнение ее надежность в реальной клинической практике.

В дальнейшем подход к обучению был изменен. Нами было принято решение учитывать не только форму, но и отдельные признаки ВМД, при этом отмечался не просто факт наличия того или иного признака («есть/нет»), а то, насколько эксперты согласованы между собой. Вместо жестких меток «0» или «1» модель получала дробные значения (например, 0,7, если 7 из 10 врачей увидели признак). Такой подход помог мягче обрабатывать спорные и неоднозначные случаи, снижая влияние разногласий между экспертами.

Обучение проводилось на тех же нейросетях – ConvNeXtV2 Base и Large. Главный результат – модель стала работать точнее (вырос F1-score), а на визуализации Grad-CAM оказалось, что новая модель действительно фокусируется на тех участках изображения, где видны признаки патологии. Это объясняется тем, что при жесткой разметке примеры с неопределенными признаками патологии заставляют модель искать любую отличительную информацию в изображении, чтобы поддержать жесткую метку, что приводит к активациям в нерелевантных зонах, и данный опыт закрепляется. При мягких метках вес таких примеров в обучении снижается пропорционально уверенности экспертов, и модель учится ассоциировать признак патологии с более характерными областями, а не случайными деталями изображения.

■ РЕЗУЛЬТАТЫ

В ходе исследования была оценена эффективность модели глубокого обучения ConvNeXtV2 в распознавании ключевых признаков глазных заболеваний на изображениях сетчатки. Для интерпретации результатов soft labels (признаки ВМД), отражающие долю экспертов, согласных с наличием признака, бинаризовались (превращались в двоичные значения, т. е. в «0» или «1») с использованием порога 0,5.

Полученные результаты показали высокую диагностическую точность модели по всем основным метрикам (см. таблицу).

Средние показатели по всем классам составили: чувствительность (Sensitivity) – 0,877, специфичность (Specificity) – 0,966, F1-мера – 0,880.

Метрики обученной модели
Trained model metrics

Признак	Sensitivity	Specificity	F1
Атрофия ретинального пигментного эпителия	0,860	0,962	0,825
Друзы	0,839	0,887	0,844
Изменения по типу отека	0,854	0,989	0,869
Инtrarетинальные кисты	0,913	0,994	0,924
Отслойка нейроэпителия	0,908	0,986	0,916
Отслойка ретинального пигментного эпителия	0,893	0,976	0,901

Модель продемонстрировала особенно высокую специфичность (не ниже 0,877 для всех признаков), что говорит о ее способности с высокой уверенностью исключать отсутствие признаков ВМД. Наиболее сложным оказался признак «Атрофия ретинального пигментного эпителия», где чувствительность была ниже, что может быть связано с визуальной неопределенностью и низкой выраженностью изменений на изображениях.

Переход от бинарных меток к мягким (признаки ВМД), отражающим степень согласия экспертов, позволил не только повысить количественные метрики, но и улучшить интерпретируемость модели.

Визуализация внимания с помощью Grad-CAM показала, что модель стала более точно фокусироваться на клинически значимых, патологических участках изображения, что подтверждает ее более осмысленный и устойчивый анализ по сравнению с предыдущими подходами.

Таким образом, предложенный метод эффективно справляется с детекцией широкого спектра ретинальных изменений при ВМД, сочетая высокую точность и клиническую интерпретируемость.

■ ОБСУЖДЕНИЕ

Полученные результаты демонстрируют, что использование меток с признаками ВМД, отражающими степень согласия экспертов, и последующая их бинаризация при анализе медицинских изображений позволяют построить модель, сочетающую высокую точность и клиническую интерпретируемость. Переход от жестких бинарных меток к непрерывным значениям на этапе обучения помог учесть неопределенность в оценках врачей, особенно в сложных, «серых» диагностических случаях, и снизить влияние случайных разногласий между экспертами.

Высокие значения специфичности (в среднем 0,966) свидетельствуют о том, что модель редко ошибается, «выдумывая» патологию там, где ее нет. Это особенно важно в клинике, где ложноположительные находки могут привести к выбору неверной тактики лечения. В то же время чувствительность на уровне 0,877 указывает на хорошую, но не идеальную способность модели обнаруживать все истинные случаи заболевания. Наиболее низкая чувствительность отмечена для атрофии ретинального пигментного эпителия: вероятно, из-за малой контрастности и размытости этих изменений на изображениях, а также из-за меньшей выраженности признака половина всех примеров имела оценку ниже 0,6, что затрудняет ее выявление как для врачей, так и для модели.

Особое значение имеет улучшение качества визуальной интерпретации: анализ с помощью Grad-CAM показал, что модель теперь действительно фокусируется на анатомически релевантных зонах, таких как кисты, отслойки пигментного эпителия и наличие различных типов жидкости.

Тем не менее остаются вызовы. Например, признаки с низкой межэкспертной согласованностью могут по-прежнему вносить шум в обучение. В будущем можно рассмотреть взвешенный подход к обучению, где степень влияния каждого эксперта учитывается в зависимости от его опыта или согласованности с группой.

В целом предложенный подход открывает перспективу создания более надежных и интерпретируемых систем искусственного интеллекта в медицинской визуализации, таких, которые не просто дают правильный ответ, а обосновывают его так,



как это делает врач. Это делает модель более пригодной для интеграции в клиническую практику в качестве вспомогательного инструмента диагностики.

■ ЗАКЛЮЧЕНИЕ

В данной работе представлен подход к автоматизированному анализу ОКТ сетчатки с использованием технологий глубокого обучения для диагностики возрастной макулярной дегенерации. Основное внимание уделено не просто распознаванию форм заболевания, а детекции отдельных патологических признаков, таких как друзы, отслойки, кисты и атрофия, которые имеют ключевое значение для постановки диагноза и выбора тактики лечения.

Ключевым достижением исследования стало внедрение мягких меток признаков ВМД, отражающих степень согласия врачей-экспертов при аннотировании изображений. В отличие от традиционного подхода с жесткими бинарными метками, использование непрерывных значений позволило модели учитывать диагностическую неопределенность и снизить влияние межэкспертного разброса. Это привело к заметному улучшению как количественных показателей (средний F1-score составил 0,880, специфичность – 0,966), так и качества интерпретации решений модели, подтвержденного визуализацией Grad-CAM: теперь нейросеть фокусируется на клинически значимых зонах, а не на фоновых артефактах.

Полученные результаты свидетельствуют о том, что правильно организованный процесс сбора и подготовки данных, включая многократную экспертизу, учет степени уверенности и структурированное аннотирование, играет не меньшую роль, чем выбор архитектуры модели. Предложенный подход способствует созданию не только точных, но и интерпретируемых систем искусственного интеллекта, которые могут быть доверены в клинической практике.

Тем не менее остаются задачи для дальнейшего развития: повышение чувствительности к слабо выраженным признакам (например, атрофии ПЭ), стандартизация данных с разных ОКТ-устройств и масштабирование модели на более разнообразные популяции пациентов. В перспективе такие системы могут стать неотъемлемой частью диагностического процесса, помогая врачам быстрее и точнее выявлять патологии, снижать нагрузку на специалистов и улучшать результаты лечения пациентов.

Таким образом, интеграция глубокого обучения с клинически осмысленным подходом к разметке данных открывает новые возможности для развития компьютерной диагностики в офтальмологии и демонстрирует, что путь к надежному ИИ в медицине лежит не только через алгоритмы, но и через тесное взаимодействие с врачебным опытом.

■ ЛИТЕРАТУРА/REFERENCES

1. Röhlig M., Rosenthal P., Schmidt C., Schumann H., Stachs O. Visual Analysis of Optical Coherence Tomography Data in Ophthalmology. *Eurova*. 2017;37–41. doi: 10.2312/EUROVA.20171117
2. Sigal I.A., Schuman J.S., Ishikawa H., Kagemann L., Wollstein G. A Problem of Proportions in OCT-Based Morphometry and a Proposed Solution. *Investigative Ophthalmology. Visual Science*. 2016;57(2):484–485. doi: 10.1167/IOVS.15-18570
3. Li D., Ran A., Cheung C.Y., Prince J.L. Deep learning in optical coherence tomography: Where are the gaps? *Clinical and Experimental Ophthalmology*. 2023. doi: 10.1111/ceo.14258
4. Yanagihara R.T., Lee C.S., Ting D.S.W., Lee A.Y. Methodological Challenges of Deep Learning in Optical Coherence Tomography for Retinal Diseases: A Review. *Translational Vision Science. Technology*. 2020;9(2):11. doi: 10.1167/TVST.9.2.11
5. Lang A. *Methods for automated analysis of macular OCT data*. 2016. Available at: <https://jscholarship.library.jhu.edu/handle/1774.2/58721>.