

DOI: <https://doi.org/10.17816/DD623801>

Классификация снимков оптической когерентной томографии с использованием методов глубокого машинного обучения

А.А. Арзамасцев^{1,2}, О.Л. Фабрикантов², Е.В. Кулагина², Н.А. Зенкова³¹ Воронежский государственный университет, Воронеж, Россия;² Национальный медицинский исследовательский центр «Межотраслевой научно-технический комплекс «Микрохирургия глаза» имени академика С.Н. Фёдорова», Тамбов, Россия;³ Тамбовский государственный университет имени Г.Р. Державина, Тамбов, Россия

АННОТАЦИЯ

Обоснование. Оптическая когерентная томография — современный высокотехнологичный и информативный метод выявления патологии сетчатки глаза и преретинальных слоёв стекловидного тела. Однако описание и интерпретация результатов исследования требуют высокой квалификации и специальной подготовки врача-офтальмолога, а также значительных временных затрат врача и пациента. Вместе с тем использование математических моделей на основе аппарата искусственных нейронных сетей в настоящее время позволяет автоматизировать многие процессы, связанные с обработкой изображений. Именно поэтому актуально решение задач, связанных с автоматизацией процесса классификации снимков оптической когерентной томографии на основе глубокого обучения моделей искусственных нейронных сетей.

Цель — разработать архитектуры математических (компьютерных) моделей на основе глубокого обучения свёрточных нейронных сетей, предназначенных для классификации снимков оптической когерентной томографии сетчатки глаза; сравнить результаты вычислительных экспериментов, проведённых с использованием средств Python в Google Colaboratory при одно- и многомодельном подходе, и выполнить оценки точности классификации; сделать выводы об оптимальной архитектуре моделей искусственных нейронных сетей и значениях используемых гиперпараметров.

Материалы и методы. Исходный датасет, представляющий собой обезличенные снимки оптической когерентной томографии реальных пациентов, включал более 2000 изображений, полученных непосредственно с прибора в разрешении 1920×969×24 BPP. Количество классов изображений — 12. Для создания обучающего и валидационного наборов данных осуществляли «вырезание» предметной области 1100×550×24 BPP. Изучали различные подходы: возможность использования предобученных свёрточных нейронных сетей с переносом обучения, методики изменения размера и аугментации изображений, а также различные сочетания гиперпараметров моделей искусственных нейронных сетей. При компиляции модели использовали следующие параметры: оптимизатор Adam, функцию потерь categorical_crossentropy, метрику accuracy. Все технологические процессы с изображениями и моделями искусственных нейронных сетей проводили с использованием средств языка Python в Google Colaboratory.

Результаты. Предложены одно- и многомодельный принципы классификации изображений оптической когерентной томографии сетчатки глаза. Вычислительные эксперименты по автоматизированной классификации таких изображений, полученных с томографа DRI OCT Triton, с использованием различных архитектур моделей искусственных нейронных сетей показали точность при обучении и валидации 98–100%, и на дополнительном тесте — 85%, что является удовлетворительным результатом. Выбрана оптимальная архитектура модели искусственной нейронной сети — 6-слойная свёрточная сеть, — и определены значения её гиперпараметров.

Заключение. Результаты глубокого обучения моделей свёрточных нейронных сетей с различной архитектурой, их валидации и тестирования показали удовлетворительную точность классификации снимков оптической когерентной томографии сетчатки глаза. Данные разработки могут быть использованы в системах поддержки принятия решений в области офтальмологии.

Ключевые слова: искусственный интеллект; медицинские данные; датасет; машинное обучение; свёрточные нейронные сети; оптическая когерентная томография.

Как цитировать:

Арзамасцев А.А., Фабрикантов О.Л., Кулагина Е.В., Зенкова Н.А. Классификация снимков оптической когерентной томографии с использованием методов глубокого машинного обучения // Digital Diagnostics. 2024. Т. 5, № 1. С. 5–16. DOI: <https://doi.org/10.17816/DD623801>

Рукопись получена: 24.11.2023

Рукопись одобрена: 12.02.2024

Опубликована online: 11.03.2024



DOI: <https://doi.org/10.17816/DD623801>

Classification of optical coherence tomography images using deep machine-learning methods

Alexander A. Arzamastsev^{1,2}, Oleg L. Fabrikantov², Elena V. Kulagina², Natalia A. Zenkova³

¹ Voronezh State University, Voronezh, Russia;

² The S. Fyodorov Eye Microsurgery Federal State Institution, Tambov, Russia;

³ Derzhavin Tambov State University, Tambov, Russia

ABSTRACT

BACKGROUND: Optical coherence tomography is a modern high-tech, insightful approach to detecting pathologies of the retina and preretinal layers of the vitreous body. However, the description and interpretation of study findings require advanced qualifications and special training of ophthalmologists and are highly time-consuming for both the doctor and the patient. Moreover, mathematical models based on artificial neural networks now allow for the automation of many image processing tasks. Therefore, addressing the issues of automated classification of optical coherence tomography images using deep learning artificial neural network models is crucial.

AIM: To develop architectures of mathematical (computer) models based on deep learning of convolutional neural networks for the classification of retinal optical coherence tomography images; to compare the results of computational experiments conducted using Python tools in Google Colaboratory with single-model and multimodel approaches, and evaluate classification accuracy; and to determine the optimal architecture of models based on artificial neural networks, as well as the values of the hyperparameters used.

MATERIALS AND METHODS: The original dataset included >2,000 anonymized optical coherence tomography images of real patients, obtained directly from the device with a resolution of 1,920×969×24 BPP. The number of image classes was 12. To create the training and validation datasets, a subject area of 1,100×550×24 BPP was “cut out”. Various approaches were studied: the possibility of using pretrained convolutional neural networks with transfer learning, techniques for resizing and augmenting images, and various combinations of the hyperparameters of models based on artificial neural networks. When compiling a model, the following parameters were used: Adam optimizer, categorical_crossentropy loss function, and accuracy. All technological operations involving images and models based on artificial neural networks were performed using Python language tools in Google Colaboratory.

RESULTS: Single-model and multimodel approaches to the classification of retinal optical coherence tomography images were developed. Computational experiments on the automated classification of such images obtained from a DRI OCT Triton tomograph using various architectures of models based on artificial neural networks showed an accuracy of 98–100% during training and validation, and 85% during an additional test, which is a satisfactory result. The optimal architecture of the model based on an artificial neural network, a six-layer convolutional network, was selected, and the values of its hyperparameters were determined.

CONCLUSION: Deep training of convolutional neural network models with various architectures, as well as their validation and testing, resulted in satisfactory classification accuracy of retinal optical coherence tomography images. These findings can be used in decision support systems in ophthalmology.

Keywords: artificial intelligence; medical data; dataset; machine learning; convolutional neural networks; optical coherence tomography.

To cite this article:

Arzamastsev AA, Fabrikantov OL, Kulagina EV, Zenkova NA. Classification of optical coherence tomography images using deep machine-learning methods. *Digital Diagnostics*. 2024;5(1):5–16. DOI: <https://doi.org/10.17816/DD623801>

Submitted: 24.11.2023

Accepted: 12.02.2024

Published online: 11.03.2024

DOI: <https://doi.org/10.17816/DD623801>

利用深度机器学习技术对光学相干断层扫描图片进行分类

Alexander A. Arzamastsev^{1,2}, Oleg L. Fabrikantov², Elena V. Kulagina², Natalia A. Zenkova³¹ Voronezh State University, Voronezh, Russia;² The S. Fyodorov Eye Microsurgery Federal State Institution, Tambov, Russia;³ Derzhavin Tambov State University, Tambov, Russia

摘要

论证。光学相干断层扫描是一种现代高科技和信息丰富的方法，用于检测视网膜和玻璃体视网膜前层的病理。然而，对研究结果的描述和解释需要眼科医生的高技能和专业培训，还需要医生和病人花费大量时间。与此同时，如今使用基于人工神经网络设备的数学模型可以实现许多图片处理过程的自动化。因此，在人工神经网络模型深度学习的基础上解决光学相干断层扫描图片分类过程的自动化问题具有现实意义。

目的。本研究的目的是开发基于卷积神经网络深度学习的数学（计算机）模型结构，此类网络设计被用于视网膜光学相干断层扫描的图像分类；使用Google Colaboratory的Python工具，比较单一模型和多模型方法的计算实验结果，并评估分类准确性；就人工神经网络模型的最佳架构和所使用的超参数值得出结论。

材料与amp;方法。原始数据集是真正患者的匿名光学相干断层扫描图像，其中包括2000多张图像。图像直接从设备中获取，分辨率为1920×969×24 BPP。图像类别数量为12个。为了创建训练和验证数据集，对1100×550×24 BPP主题区域进行了“切割”。研究了不同的方法：使用带学习转移的预训练卷积神经网络的可能性、图像缩放和增强技术，以及人工神经网络模型超参数的不同组合。在编制模型时，使用了以下参数：Adam优化器、categorical_crossentropy损失函数、accuracy指标。所有图像和人工神经网络模型的技术处理都是通过Google Colaboratory的Python语言工具进行的。

结果。提出了视网膜光学相干断层扫描图片分类的单模型和多模型原理。使用不同的人工神经网络模型架构对来自DRI OCT Triton断层扫描仪的此类图像进行自动分类的计算实验表明了，训练和验证期间的准确率为98%–100%，额外测试的准确率为85%。这是一个令人满意的结果。人工神经网络模型的最佳架构是6层卷积网络，并确定了其超参数值。

结论。对不同结构的卷积神经网络模型进行深度学习、验证和测试的结果表明，视网膜光学相干断层扫描图片分类的准确度令人满意。这些研制成果可被用于眼科领域的决策支持系统。

关键词：人工智能；医学数据；数据集；机器学习；卷积神经网络；光学相干断层扫描。

引用本文：

Arzamastsev AA, Fabrikantov OL, Kulagina EV, Zenkova NA. 利用深度机器学习技术对光学相干断层扫描图片进行分类. *Digital Diagnostics*. 2024;5(1):5–16. DOI: <https://doi.org/10.17816/DD623801>

收到: 24.11.2023

接受: 12.02.2024

发布日期: 11.03.2024

ОБОСНОВАНИЕ

Оптическая когерентная томография (ОКТ) — современный высокотехнологичный и информативный метод выявления патологии сетчатки глаза и преретинальных слоёв стекловидного тела [1]. Однако описание и интерпретация результатов исследования требуют высокой квалификации и специальной подготовки врача, а также значительных временных затрат офтальмолога и пациента. По этой причине решение задач, связанных с автоматизацией процесса классификации снимков ОКТ, является актуальным.

Вместе с тем в настоящее время наблюдается быстрое развитие компьютерного инструментария и технологий, позволяющих создавать системы искусственного интеллекта на основе нейронных сетей различной архитектуры, как медицинских [2, 3], так и общего назначения [4–6]. Современные офтальмологические центры за последние десятилетия создали хранилища данных о пациентах, включающие десятки и сотни тысяч снимков ОКТ, что открывает путь к поиску обобщённых зависимостей и связей между отдельными параметрами и построению принципиально новых подходов для идентификации, классификации, расчётов и предсказаний на основе научного подхода, ядром которого практически всегда является математическая модель.

В одной из своих работ мы уже описали компьютерные методы анализа стекловидного тела, выделения и аппроксимации границы сетчатки, определения кривизны границы, расчётов средней толщины сетчатки и др., выполненных в том числе и с использованием искусственных нейронных сетей (ИНС) [7]. В данной статье, являющейся логическим продолжением этой работы, представлены результаты, полученные при классификации снимков ОКТ с использованием свёрточных нейронных сетей (СНС) с использованием одно- и многомодельного подходов.

Необходимо отметить, что решению близких проблем посвящено несколько опубликованных работ. Так, Ю.А. Васильев и соавт. [8] разработали общую методологию тестирования и мониторинга программного обеспечения на основе технологий искусственного интеллекта для медицинской диагностики. Методология направлена на повышение качества данного программного обеспечения и его внедрение в практическое здравоохранение. Она состоит из семи этапов: самотестирование, функциональное тестирование, калибровочное тестирование, технологический мониторинг, клинический мониторинг, обратная связь и доработка. Отличительной особенностью методологии является цикличность этапов тестирования, мониторинга и доработки программного обеспечения, приводящая к постоянному повышению его качества, а также наличие подробных требований к результатам его работы и участие врачей в его оценке. Методология позволяет разработчикам программного обеспечения

достичь высоких результатов и продемонстрировать достижения в различных направлениях, а пользователям — сделать осознанный и уверенный выбор среди программ, прошедших независимую и всестороннюю проверку качества.

Целью статьи Е.А. Каталевской и соавт. [9] являлась разработка алгоритмов сегментации визуальных признаков диабетической ретинопатии и диабетического макулярного отёка на цифровых фотографиях глазного дна, сделанных с помощью фундус-камеры. Для сегментации были выбраны признаки, включённые в Международную классификацию: микроаневризмы, твёрдые экссудаты, мягкие экссудаты, интратретинальные геморрагии, неоваскуляризация сетчатки и диска зрительного нерва, преретинальные геморрагии, эпиретинальный фиброз, лазерные коагуляты. Реализацию нейронных сетей и процесс их обучения осуществляли с помощью фреймворка глубокого обучения TensorFlow (Google Brain, США). Объём обучающей базы данных составил 1200 фотографий, в то время как для валидации использовали 310 фотографий глазного дна. Точность определения указанных признаков по обученной модели составила от 86% до 96%.

В работе Т. Керр и соавт. [10] описана система домашнего мониторинга возрастной макулярной дегенерации. СНС здесь использовали для сегментации всей сетчатки, а также отслоений пигментного эпителия. Весь датасет, включающий 711 изображений, был разделён на части: обучение/валидация/тест в пропорции 60%:20%:20%. Показано, что подход на основе СНС позволяет сегментировать сетчатку достаточно точно.

Целью работы С.Н. Сахнова и соавт. [11] была разработка модели скрининга катаракты на основе открытого набора данных, а также её валидация на клинических данных. При этом открытый набор данных состоял из 9668 изображений, полученных с помощью камеры смартфона, из которых 4514 изображений относились к классу «катаракта», а 5154 — к здоровым глазам. Набор для внешней валидации содержал 51 изображение с катарактой и нормой. Для создания модели машинного обучения использовали СНС. Значение точности классификации данных составило 97% для внутреннего валидационного набора, и 75% — для внешнего. Прогностическое значение, по мнению авторов, было низким, и был сделан вывод о том, что необходимо произвести доработку модели и обеспечить необходимые уровни метрик эффективности.

В работе С.В. Шухаева и соавт. [12] использовали предобученные сети ResNet-18, ResNet-50, VGG16, VGG19 и GoogleNet для решения задачи применимости СНС для автоматического определения дистрофии Фукса. Была произведена выборка случайных ($n=700$) биомикроскопических снимков роговицы, полученных при помощи эндотелиального микроскопа Tomey EM-3000 (Tomey Corporation, Япония). На первом этапе выполнялось разделение снимков на 2 группы. Первая группа включала снимки с наличием дистрофии Фукса, вторая — норму

или другую патологию. Снимки плотности эндотелиальных клеток были разделены на три категории: тренировочный, валидационный и тестовый наборы данных. В результате апробации нейронной сети на тестовой выборке были получены следующие значения F-метрики для различных архитектур СНС: ResNet-18 — 0,985; ResNet-50 — 1,000; VGG16 — 0,940; VGG19 — 0,990; GoogleNet — 0,987. ResNet-50 показала лучший результат на данных ImageNet с замороженными слоями, оптимизатором Adam и кросс-энтропией в качестве функции потерь.

Таким образом, краткий анализ указанных работ позволяет сделать выводы о перспективности использования ИНС-моделей на основе СНС для решения задачи классификации изображений ОКТ сетчатки глаза.

ЦЕЛЬ

Разработать архитектуры математических (компьютерных) моделей на основе глубокого обучения СНС для классификации снимков ОКТ с использованием библиотек Python Keras и Tensorflow в Google Colaboratory. Сравнить результаты вычислительных экспериментов по классификации снимков ОКТ, полученных при одно- и многомодельном подходе и выполнить оценки точности такой классификации. Сделать выводы об оптимальной с точки зрения точности классификации архитектуре ИНС-моделей и значениях используемых гиперпараметров.

МАТЕРИАЛЫ И МЕТОДЫ

Первоначальный датасет представлял собой обезличенные снимки ОКТ реальных пациентов и включал 1004 изображения, полученных непосредственно с томографа DRI OCT Triton (Topcon Corporation, Япония) в разрешении 1920×969×24 BPP в виде файлов формата JPG. В соответствии с целями классификации весь датасет был разделён опытными врачами-офтальмологами на 12 классов:

- 1 — норма;
- 2 — кистозный макулярный отёк;
- 3 — отслойка нейроэпителия;
- 4 — отслойка пигментного эпителия;
- 5 — твёрдые экссудаты;
- 6 — эпиретинальный фиброз;
- 7 — витреомакулярная адгезия;
- 8 — задняя отслойка стекловидного тела;
- 9 — сквозной макулярный разрыв + эпиретинальный фиброз;
- 10 — твёрдые экссудаты + кистозный макулярный отёк;
- 11 — друзы пигментного эпителия;
- 12 — ламеллярный разрыв + эпиретинальный фиброз.

Количество изображений каждого класса соответствовало частоте встречаемости соответствующей патологии у пациентов. В последующих вычислительных

экспериментах датасет был дополнен новыми снимками ОКТ, так что их общее число составило более 2000 изображений. Для создания обучающего, валидационного и тестового наборов данных осуществляли «вырезание» из всего снимка предметной области 1100×550×24 BPP. Весь датасет при проведении вычислительных экспериментов, как правило, был разделён на три части: обучение/валидация/тест, в пропорции 70%:20%:10%.

Были использованы также следующие технологические приёмы:

- изменение размера изображения (rescale) с использованием фильтров NEAREST, BILINEAR, BICUBIC, LANCZOS;
- аугментация данных (augmentation) с использованием различных опций — поворот изображения на заданный угол, смещение изображения по направлениям X и Y, горизонтальный и вертикальный повороты, изменение яркости канала изображения.

Отметим, что при использовании фильтров наилучшие результаты в нашем случае показал простейший NEAREST, учитывающий параметры ближайшего пикселя. Более сложные фильтры, осуществляющие аппроксимацию области различными методами, показали худшие результаты. По всей видимости, это происходило по причине того, что при сглаживании терялись мелкие детали изображений, важные для классификации.

При компиляции модели использовали следующие параметры:

- оптимизатор Adam — один из самых эффективных алгоритмов оптимизации;
- функция потерь categorical_crossentropy — категориальная перекрёстная энтропия;
- метрика accuracy — доля правильных ответов алгоритма.

Необходимо отметить, что метрика accuracy обычно используется для решения задачи классификации, если группы являются сбалансированными по количеству снимков. В нашем случае из-за малого числа снимков в обучающей и тестовой выборках было решено сделать общую среднюю оценку.

Все технологические процессы с моделями проводили с использованием средств языка Python в Google Colaboratory.

РЕЗУЛЬТАТЫ

Предварительные вычислительные эксперименты

В предварительных вычислительных экспериментах оценивали эффективность для классификации снимков ОКТ различных подходов (таких как возможность использования предобученных сетей и переноса обучения), методик изменения размера и аугментации изображений, а также сочетаний гиперпараметров ИНС-моделей (число свёрточных и полносвязных слоёв, размер пакета и т.д.).

Для предобученных нейронных сетей на основе MobileNetV2 и MobileNetV3 на обучающей выборке достигнута точность 95–98%, на валидации — 61–80%, на тесте — 41–59%. При этом использовали инструмент для масштабирования изображения до размера 224×224 пикселей, чтобы добиться соответствия с MobileNet.

Выполнено также сравнение обучаемости на указанном датасете различных предобученных нейронных сетей с переносом обучения. Получены следующие результаты их валидации: для MobileNetV2 — 80%, для ResNet101V2 — 81%, для InceptionResNetV2 — 79%, для NASNetLarge — 80%.

Для многослойных СНС, имеющих традиционную архитектуру (несколько свёрточных слоёв Conv2D, в каждом из которых имеется и функция подвыборки MaxPooling2D, преобразование массивов в одномерный тензор Flatten и несколько полносвязных слоёв Dense), на обучающей выборке достигается точность в 70–100% при разумном выборе количества эпох. Однако на валидационной выборке этот показатель значительно ниже и имеет больший разброс (35–94%). В двух случаях точность валидации оказалась выше точности, показанной на обучении, что можно связать со значительной неоднородностью данных обучающей и валидационной выборок. На тестах точность оказалась ещё ниже и составила всего 27–59%, что, конечно, не могло быть признано удовлетворительным результатом.

В результате предварительных экспериментов были сделаны следующие выводы:

- обучающая выборка является несбалансированной и неоднородной и нуждается в коррекции и дополнении новыми снимками;

- хотя модели с переносом обучения показали несколько лучшие результаты классификации, эти результаты всё ещё недостаточны для их использования в офтальмологической практике, а за счёт «заморозки» первых скрытых слоёв остаётся не очень много возможностей для их усовершенствования;
- для достижения приемлемой точности классификации требуется оптимизация гиперпараметров и, возможно, самого подхода к классификации.

Вычислительные эксперименты: одномодельный подход

С учётом результатов, полученных в ходе предварительных экспериментов, датасет был дополнен новыми снимками ОКТ, так что общее число изображений стало более 2000. В ходе экспериментов тестировали различные архитектуры многослойных последовательных (sequential) СНС: несколько свёрточных слоёв Conv2D, на выходе каждого из которых имелся слой подвыборки MaxPooling2D, а также слой, осуществляющий преобразование в одномерный тензор Flatten, и два полносвязных слоя Dense, последний из которых имел передаточную функцию нейронов softmax, удобную для решения задачи классификации.

Для разрешения нормализованного снимка 512×512 пикселей, к которому приводили все изображения датасета с использованием инструмента rescale, предельное число свёрточных слоёв составляет 7. Тестировали структуры СНС с числом таких слоёв от 2 до 7 (табл. 1), одновременно подбирая размер и количество фильтров в слоях. Для всех вычислительных экспериментов проводили обучение (как правило, с числом эпох epochs=15,

Таблица 1. Сравнение различных последовательных моделей искусственных нейронных сетей

Число свёрточных слоёв	Число оптимизируемых параметров	Точность обучения, %	Точность валидации, %	Точность на дополнительном тесте, %	Примечание
2	31 844 921	13	0	0	ИНС-модель плохо обучается
3	13 401 045	97	100	55	Число эпох обучения ИНС-модели: больше 15. Количество фильтров в слоях СНС: 3/8/16
3	15 215 889	100	100	62	Число эпох обучения ИНС-модели: 9. Количество фильтров в слоях СНС: 4/8/16
3	13 401 933	100	100	64	Число эпох обучения ИНС-модели: 12. Количество фильтров в слоях СНС: 5/8/16
4	6 929 729	100	100	65	Число эпох обучения ИНС-модели: 12
5	1 430 977	100	100	70	Число эпох обучения ИНС-модели: 9
6	556 673	100	100	85	Число эпох обучения ИНС-модели: 15
7	132 801	100	89	74	Число эпох обучения ИНС-модели: больше 15
8	—	—	—	—	8 слоёв СНС невозможно использовать для принятого размера изображений

Примечание. ИНС — искусственная нейронная сеть; СНС — свёрточная нейронная сеть.

размером пакета `BATCH_SIZE=50` и параметрами компиляции `optimizer='adam'`, `loss='categorical_crossentropy'`, `metrics=['accuracy']`, а также валидацию и дополнительное тестирование на изображениях, ранее не включённых в датасет.

Практически для всех ИНС-моделей, за исключением двухслойной, достигнута приемлемая точность обучения и валидации (см. табл. 1). Однако если сравнить точность прохождения ими дополнительного теста, то она сначала возрастает по мере увеличения числа свёрточных слоёв, достигая максимального значения 85% в шестислойной модели, а затем убывает для семислойной. Необходимо отметить, что в настоящее время приведённые здесь модели позиционируются исключительно как система поддержки принятия решения врача-офтальмолога. Учитывая, что датасет содержит лишь ограниченное число снимков различных патологий, уровень 85% был принят нами как достаточный и соответствующий классификации снимков врачом-офтальмологом с небольшим опытом работы. Сделан вывод о том,

что существует оптимальное, с точки зрения точности, число слоёв, которое в данном случае равно 6. Однако это значение может быть в дальнейшем изменено по мере накопления новых данных в датасете и переобучения моделей.

На рис. 1 показаны процессы обучения и валидации ИНС-моделей с числом свёрточных слоёв от 4 до 7. При использовании 4 и 5 свёрточных слоёв процесс обучения завершается после 9 эпох, достигая точности на обучении и валидации 100%. При этом точность прохождения дополнительного теста составляет всего 65–70% (см. табл. 1). В случае использования модели с 6 свёрточными слоями процесс обучения занимает более длительное время — 15 эпох, также достигая точности на обучении и валидации 100%. Однако точность прохождения дополнительного теста возрастает до 85%, что сочтено нами за удовлетворительный результат. При дальнейшем увеличении числа свёрточных слоёв до 7 процесс обучения и валидации ИНС-модели занимает более 15 эпох, при этом точность обучения составляет 100%, а точность

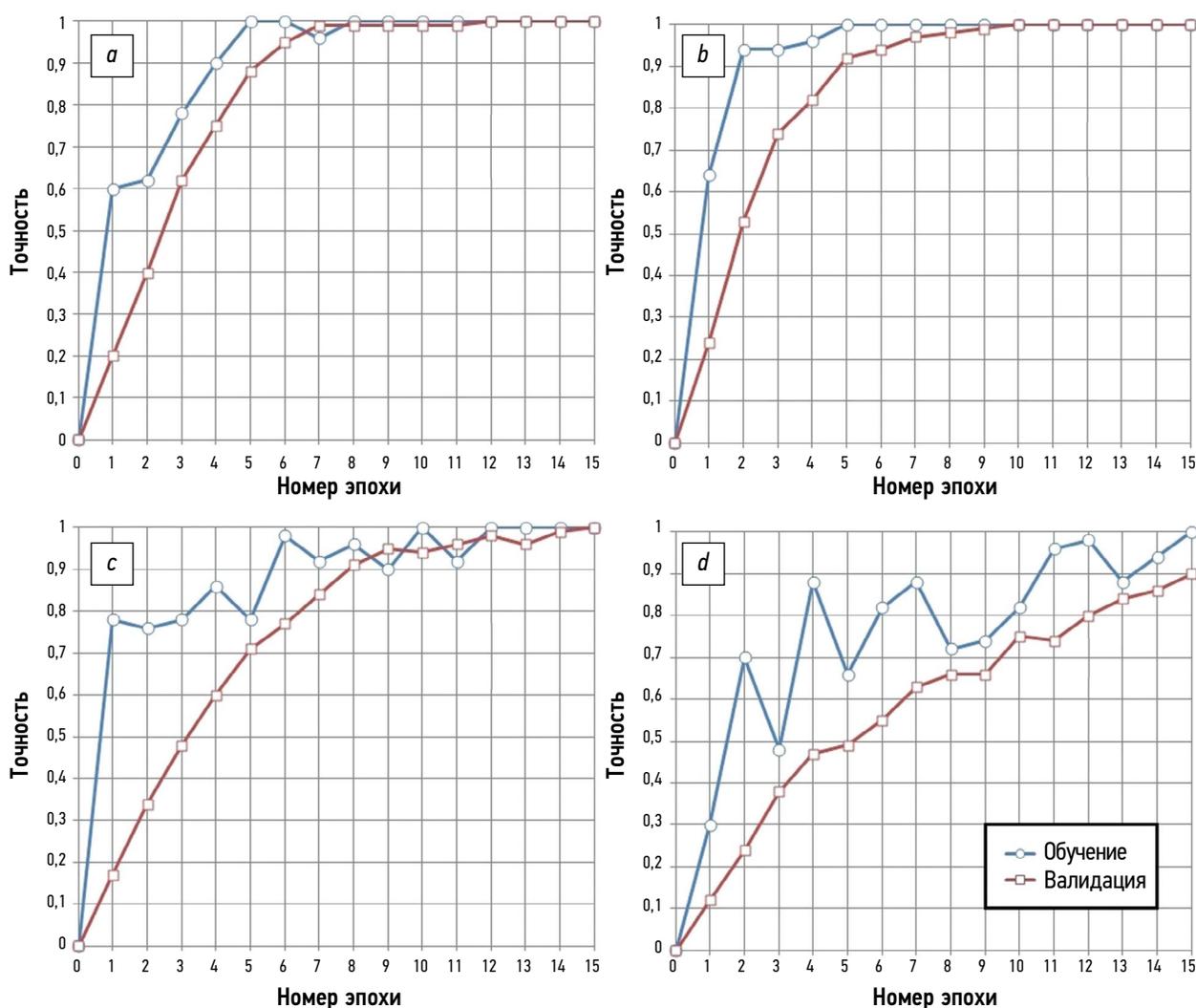


Рис. 1. Обучение и валидация моделей свёрточных нейронных сетей: *a* — четыре свёрточных слоя; *b* — пять свёрточных слоёв; *c* — шесть свёрточных слоёв; *d* — семь свёрточных слоёв.

```
[ ] model = tf.keras.models.Sequential([
    tf.keras.layers.Conv2D(4, (9,9), activation='relu', input_shape=(IMG_SHAPE, IMG_SHAPE, 3)),
    tf.keras.layers.MaxPooling2D(2, 2),

    tf.keras.layers.Conv2D(8, (5, 5), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),

    tf.keras.layers.Conv2D(16, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),

    tf.keras.layers.Conv2D(32, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),

    tf.keras.layers.Conv2D(32, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),

    tf.keras.layers.Conv2D(64, (3, 3), activation='relu'),
    tf.keras.layers.MaxPooling2D(2, 2),

    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(225, activation='relu'),
    tf.keras.layers.Dense(12, activation='softmax')
])
```

Рис. 2. Архитектура и параметры модели искусственной нейронной сети с шестью свёрточными слоями. Первая цифра в Conv2D — число используемых фильтров в слое свёртки, две следующие цифры — размер фильтра в пикселях. Активационные функции нейронов сети — relu, в выходном слое классификации — softmax. Первая цифра в полносвязном слое Dense — число нейронов.

валидации и прохождения теста снижается до 89% и 74% соответственно (см. рис. 1 и табл. 1).

На рис. 2 показана архитектура оптимальной ИНС-модели с точки зрения точности классификации снимков ОКТ сетчатки глаза. Она включает 6 свёрточных слоёв Conv2D с подвыборкой MaxPooling2D, слой Flatten и два полносвязных слоя Dense, выполняющих роль классификатора, последний из которых имеет передаточную функцию нейронов softmax.

Объяснением более низкой (по сравнению с обучением и валидацией) точности классификации, полученной на тестах, является следующее обстоятельство. Для обучения моделей используются относительно небольшие датасеты, содержащие порядка 2000 снимков, которые, по всей видимости, не содержат полный набор графических деталей, характерных для той или иной патологии. Если в тестовой выборке встречаются такие детали, то классификация может быть неверной даже при условии 100%-й точности, полученной при валидации ИНС-модели.

Предварительные вычислительные эксперименты: многомодельный подход

Следуя общей логике исследования, а также учитывая необходимость повышения точности прохождения дополнительного теста, О.Л. Фабрикантовым и Е.В. Кулагиной была предложена последовательная схема, в общих чертах имитирующая процесс идентификации изображения ОКТ сетчатки глаза врачом-офтальмологом. На её основе был построен компьютерный алгоритм (см. рис. 2).

Данный алгоритм предполагает многомодельный последовательный принцип реализации (рис. 3).

На первом этапе производится предварительная обработка снимков (блоки 1–3). В блоке 4 используется ИНС-модель 1, которая осуществляет предварительную классификацию — отличает норму от какой-либо патологии. Результат такой классификации сохраняется (S1). Такая модель должна быть обучена и подвергнута валидации на специальном датасете 1, содержащем лишь два соответствующих класса изображений. Если патология не обнаружена (блок 5), то осуществляем переход к четвёртому этапу — анализу стекловидного тела, минуя все промежуточные этапы. Аналогично в блоке 6: используем ИНС-модель 2, обученную на специальном датасете 2 и позволяющую выявить, имеется ли макулярный разрыв, или он отсутствует. Результат сохраняется (S2). Если макулярный разрыв есть (блок 7), то в блоке 8 на основе ИНС-модели 3, также обученной на специальном датасете 3, определяется, имеет ли место сквозной или ламеллярный макулярный разрыв, после чего сохраняются результаты (S3), и осуществляется переход ко второму этапу.

Если макулярного разрыва нет (блок 7), то осуществляется переход к блоку 9, в котором с помощью ИНС-модели 4, обученной на специальном датасете 4, выявляется один из трёх вариантов — кистозный макулярный отёк, диффузный макулярный отёк или их отсутствие. Результат сохраняется (S4), и осуществляется переход ко второму этапу.

На втором этапе с помощью ИНС-модели 5 (блок 10), обученной на специальном датасете 5, выявляется один из трёх вариантов — отслойка нейрозпителя, отслойка пигментного эпителия или отсутствие указанных

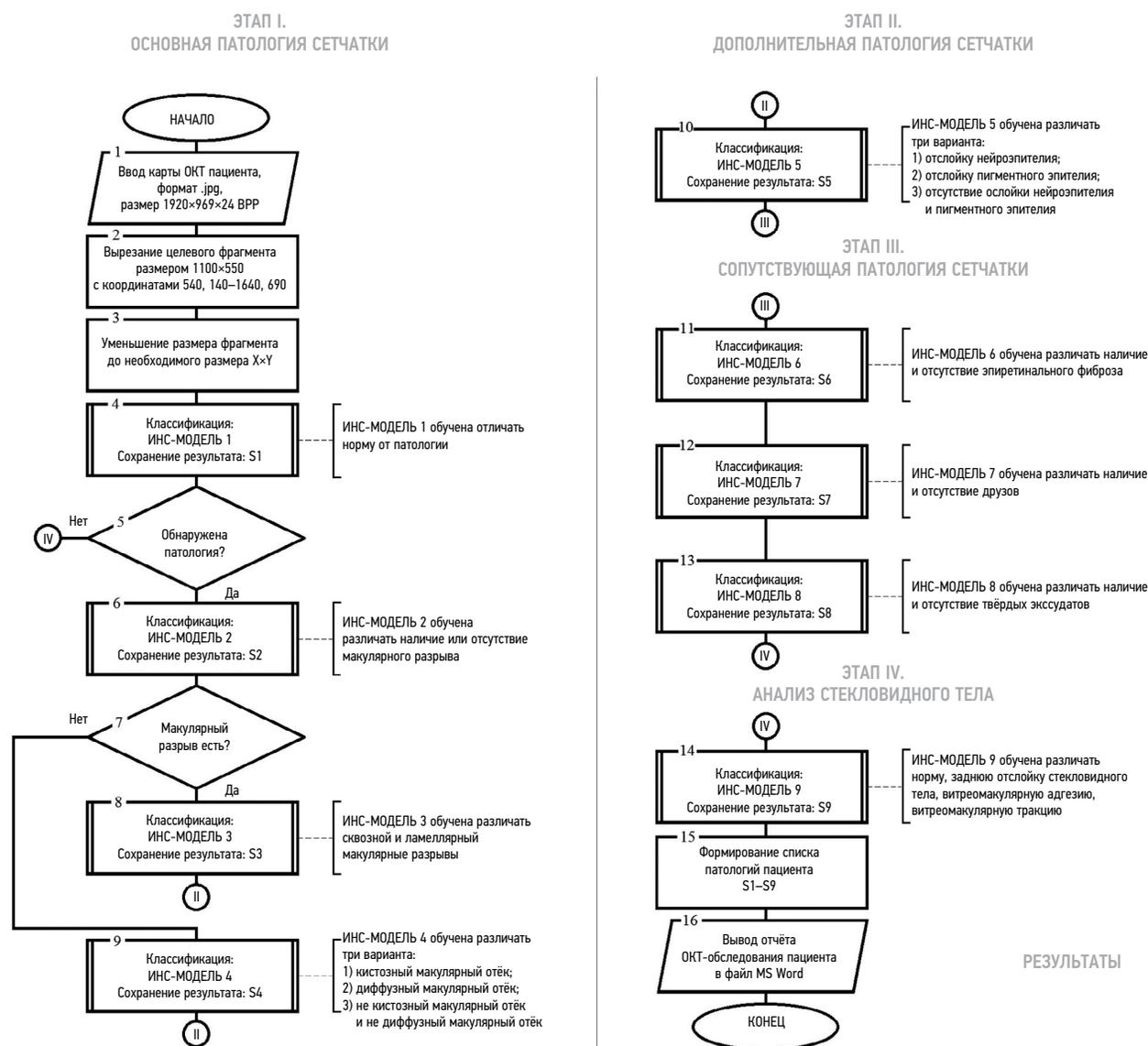


Рис. 3. Блок-схема многомодельного алгоритма идентификации снимков оптической когерентной томографии. ИНС — искусственная нейронная сеть; ОКТ — оптическая когерентная томография.

вариантов, после чего осуществляется сохранение результата (S5) и переход к третьему этапу анализа и классификации снимков ОКТ.

На третьем этапе происходит последовательное использование ИНС-моделей 6, 7 и 8.

- ИНС-модель 6 (блок 11) обучена на специальном датасете 6 распознаванию наличия или отсутствия эпиретинального фиброза.
- ИНС-модель 7 (блок 12) обучена на специальном датасете 7 распознаванию наличия или отсутствия друзов.
- ИНС-модель 8 (блок 13) обучена на специальном датасете 8 распознаванию наличия или отсутствия эксудатов.

В блоках 11–13 осуществляется также и сохранение соответствующих результатов (S6, S7, S8), после чего выполняется переход к четвёртому этапу.

На четвёртом этапе происходит использование ИНС-модели 9 (блок 14), обученной на специальном датасете 9 распознаванию нормы, задней отслойки стекловидного тела, витреомакулярной адгезии и витреомакулярной тракции. Результаты сохраняются (S9). В блоках 15 и 16 происходит формирование общего списка патологий на основании ранее сохранённых S1–S9 и вывод отчёта с описанием карты ОКТ в файл.

ОБСУЖДЕНИЕ

В описанном подходе для классификации патологии по снимкам ОКТ используются 9 различных ИНС-моделей, каждая из которых обучена на своём уникальном датасете (1–9). Следует отметить, что на завершающем этапе анализа стекловидного тела вместо ИНС-модели 9 может

быть использован алгоритм, описанный нами ранее [7]. Он включает:

- вертикальное сканирование снимка и определение координат X и Y границ стекловидного тела;
- сглаживание координат Y по методу скользящего среднего с базой, соизмеримой с минимальными деталями изображения (в нашем случае — 10 пикселей), аппроксимацией границы стекловидного тела сплайном или параболой соответствующего порядка;
- расчёт максимальной кривизны границы и соответствующих расстояний для идентификации задней отслойки стекловидного тела, витреомакулярной адгезии и витреомакулярной тракции.

В настоящее время осуществляется тестирование многомодельного алгоритма (см. рис. 3) одновременно с увеличением количества снимков ОКТ в датасетах и оптимизацией гиперпараметров ИНС-моделей. Предварительные вычислительные эксперименты, проведённые для нескольких этапов этого алгоритма, показали, что здесь удаётся добиться точности в 98–100% на обучении и валидации, при этом точность прохождения дополнительного теста возрастает по сравнению с одномоделным подходом за счёт уменьшения числа классифицируемых на каждом этапе факторов. При этом для всех ИНС-моделей 1–9 используется единая архитектура, имеющая 7 свёрточных слоёв. Различаются они лишь тем, что обучены на уникальных датасетах и, соответственно, имеют различные наборы коэффициентов межнейронных синаптических связей.

ЗАКЛЮЧЕНИЕ

Предложены одно- и многомодельные принципы классификации изображений ОКТ сетчатки глаза. Вычислительные эксперименты по автоматизированной классификации таких изображений, полученных с томографа DRI OCT Triton, с использованием различных архитектур ИНС-моделей показали точность при обучении и валидации 100%, и на дополнительном тесте — 85%, что является удовлетворительным результатом.

Выбрана оптимальная архитектура ИНС-модели — 6-слойная СНС — и определены значения её гиперпараметров. Указанные обстоятельства позволяют использовать разработки в системах поддержки принятия решений в области офтальмологии.

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Источник финансирования. Авторы заявляют об отсутствии внешнего финансирования при проведении исследования.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Вклад авторов. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли

существенный вклад в разработку концепции, проведение исследования и подготовку статьи, прочли и одобрили финальную версию перед публикацией). Наибольший вклад распределён следующим образом: А.А. Арзамасцев — разработка концепции, предварительная обработка изображений ОКТ, проведение исследования, написание программ, проведение вычислительных экспериментов, многомодельный подход к классификации изображений, подготовка рукописи; О.Л. Фабрикантов — разработка концепции, сбор и подготовка изображений ОКТ, многостадийная схема классификации изображений ОКТ; обсуждение и утверждение итогового варианта рукописи; Е.В. Кулагина — разработка методологии сбора и подготовки изображений ОКТ, проведение исследования, многостадийная схема классификации изображений ОКТ, утверждение итогового варианта рукописи; Н.А. Зенкова — разработка концепции, проведение исследования, редактирование и утверждение итогового варианта текста рукописи, анализ литературных данных, редактирование текста статьи.

Благодарности. Работа выполнена в соответствии с договором о научно-техническом сотрудничестве Воронежского государственного университета и Федерального государственного автономного учреждения «Национальный медицинский исследовательский центр «Микрохирургия глаза» имени академика С.Н. Федорова», Тамбовского филиала от 28.11.2022.

В предварительных вычислительных экспериментах приняли участие магистранты факультета прикладной математики, информатики и механики Воронежского государственного университета, которые обучались по курсу «Системы искусственного интеллекта на базе искусственных нейронных сетей» в 2023 году в рамках лабораторного практикума: Е.П. Гализина, В.А. Гущина, И.О. Завьялова, В.Ю. Колупаев, Н.М. Кушнарёв, И.Ю. Новоскольцев, Е.А. Струкова, Н.М. Чернышов, И.Д. Чикуннов, А.А. Щеглевых, а также магистрант-дипломник М.А. Курприн. Эти работы выполнены под руководством одного из авторов статьи.

ADDITIONAL INFORMATION

Funding source. This study was not supported by any external sources of funding.

Competing interests. The authors declare that they have no competing interests.

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work.

А.А. Арзамасцев — development of the concept, preliminary processing of OCT images, conducting research, writing programs, conducting computational experiments, multi-model approach to image classification, preparation of the manuscript; О.Л. Фабрикантов — development of the concept, collection and preparation of OCT images, multi-stage classification scheme for OCT images, discussion and approval of the final version of the

manuscript; E.V. Kulagina — development of a methodology for collecting and preparing OCT images, conducting research, a multi-stage classification scheme for OCT images, approval of the final version of the manuscript; N.A. Zenkova — concept development, research, editing and approval of the final version of the manuscript, analysis of literature data, editing the text of the article.

Acknowledgments. The work was carried out in accordance with the agreement on scientific and technical cooperation between Voronezh State University and the Federal State Autonomous Institution "National Medical Research Center" Interindustry Scientific

and Technical Complex "Eye Microsurgery" named after Academician S.N. Fedorov", Tambov branch dated November 28, 2022.

Master's degree students from the Faculty of Applied Mathematics, Informatics and Mechanics of Voronezh State University took part in the preliminary computational experiments: E.P. Galizina, V.A. Gushchina, I.O. Zavyalova, V.Yu. Kolupaev, N.M. Kushnarev, I.Yu. Novoskoltsev, E.A. Strukova, N.M. Chernyshov, I.D. Chikunov, A.A. Shcheglevatykh, as well as a diploma student M.A. Kuprin. These works were carried out under the guidance of one of the authors of the article.

СПИСОК ЛИТЕРАТУРЫ

1. Оптическая когерентная томография сетчатки / под ред. Дж.С. Дакера, Н.К. Вэхид, Д.Р. Голдмана. Москва : МЕДпресс-информ, 2021.
2. Oakden-Rayner L., Palme L.J. Artificial intelligence in medicine: Validation and study design. In: Ranschart E., Morozov S., Algra P., editors. *Artificial intelligence in medical imaging*. Cham : Springer, 2019. P. 83–104. doi: 10.1007/978-3-319-94878-2_8
3. Рамсундар Б., Истман П., Уолтерс П., Панде В. Глубокое обучение в биологии и медицине. Москва : ДМК Пресс, 2020.
4. Будума Н., Локашо Н. Основы глубокого обучения. Создание алгоритмов для искусственного интеллекта следующего поколения. Москва : Манн, Иванов и Фербер, 2020.
5. Фостер Д. Генеративное глубокое обучение. Творческий потенциал нейронных сетей. Санкт-Петербург : Питер, 2020.
6. Постолит А.В. Основы искусственного интеллекта в примерах на Python. Санкт-Петербург : БХВ-Петербург, 2021.
7. Арзамасцев А.А., Фабрикантов О.Л., Зенкова Н.А., Кулагина Е.В. Разработка программного обеспечения для анализа протоколов оптической когерентной томографии сетчатки глаза и автоматизированного составления их описаний // *Современные проблемы науки и образования*. 2021. № 6. EDN: PCVMRX doi: 10.17513/spno.31208
8. Васильев Ю.А., Владимировский А.В., Омелянская О.В., и др. Методология тестирования и мониторинга программного обе-

спечения на основе технологий искусственного интеллекта для медицинской диагностики // *Digital Diagnostics*. 2023. Т. 4, № 3. С. 252–267. doi: 10.17816/DD321971

9. Каталевская Е.А., Каталевский Д.Ю., Тюриков М.И., Шайхутдинова Э.Ф., Сизов А.Ю. Алгоритм сегментации визуальных признаков диабетической ретинопатии (ДР) и диабетического макулярногo отёка (ДМО) на цифровых фотографиях глазного дна // *Российский журнал телемедицины и электронного здравоохранения*. 2021. Т. 7, № 4. С. 17–26. EDN: PPSPAL doi: 10.29188/2712-9217-2021-7-4-17-26
10. Kepp T., Sudkamp H., Burchard C., et al. Segmentation of retinal low-cost optical coherence tomography images using deep learning // *Medical Imaging 2020: Computer-Aided Diagnosis*. 2020. Vol. 11314. P. 389–396. doi: 10.48550/arXiv.2001.08480
11. Сахнов С.Н., Аксенов К.Д., Аксенова Л.Е., и др. Разработка модели скрининга катаракты с использованием открытого набора данных и алгоритмов глубокого машинного обучения // *Офтальмохирургия*. 2022. № 4S. С. 13–20. EDN: VEGPAW doi: 10.25276/0235-4160-2022-4S-13-20
12. Шухаев С.В., Мордовцева Е.А., Пустозеров Е.А., Кудлахмедов Ш.Ш. Применение сверточных нейронных сетей для определения эндотелиальной дистрофии Фукса // *Офтальмохирургия*. 2022. № 4S. С. 70–76. EDN: WEZTKV doi: 10.25276/0235-4160-2022-4S-70-76

REFERENCES

1. Daker DS, Vekhid NK, Goldman DR, editors. *Optical coherence tomography of the retina*. Moscow: MEDpress-inform; 2021. (In Russ).
2. Oakden-Rayner L, Palme LJ. Artificial intelligence in medicine: Validation and study design. In: Ranschart E, Morozov S, Algra P, editors. *Artificial intelligence in medical imaging*. Cham: Springer; 2019. P.83–104. doi: 10.1007/978-3-319-94878-2_8
3. Ramsundar B, Istman P, Uolters P, Pande V. *Deep learning in biology and medicine*. Moscow: DMK Press; 2020. (In Russ).
4. Buduma N, Lokasho N. *Foundations of deep learning. Creating Algorithms for Next Generation Artificial Intelligence*. Moscow: Mann, Ivanov i Ferber; 2020. (In Russ).
5. Foster D. *Generative deep learning. Creative potential of neural networks*. Saint Petersburg: Piter; 2020. (In Russ).
6. Postolit AV. *Fundamentals of Artificial Intelligence in Python examples*. Saint Petersburg: BKhV-Peterburg; 2021. (In Russ).
7. Arzamastsev AA, Fabrikantov OL, Zenkova NA, Kulagina EV. Software development for analysing the optical coherence tomography protocols of the retina and automatic composition of their descriptions. *Sovremennyye problemy nauki i obrazovaniya*. 2021;(6). EDN: PCVMRX doi: 10.17513/spno.31208
8. Vasiliev YA, Vlazimirsky AV, Omelyanskaya OV, et al. Methodology for testing and monitoring artificial intelligence-based software for medical diagnostics. *Digital Diagnostics*. 2023;4(3):252–267. doi: 10.17816/DD321971
9. Katalevskaya EA, Katalevsky DYU, Tyurikov MI, Shaykhutdinova EF, Sizov AYU. Algorithm for segmentation of visual signs of diabetic retinopathy (DR) and diabetic macular edema (DME) in digital fundus images. *Russian Journal of Telemedicine and e-health*. 2021;7(4):17–26. EDN: PPSPAL doi: 10.29188/2712-9217-2021-7-4-17-26

10. Kepp T, Sudkamp H, Burchard C, et al. Segmentation of retinal low-cost optical coherence tomography images using deep learning. *Medical Imaging 2020: Computer-Aided Diagnosis*. 2020;11314:389–396. doi: 10.48550/arXiv.2001.08480
11. Sakhnov SN, Axenov KD, Axenova LE, et al. Development of a cataract screening model using an open dataset and deep machine learning algorithms. *Fyodorov Journal*

- of Ophthalmic Surgery*. 2022;(4S):13–20. EDN: VEGPAW doi: 10.25276/0235-4160-2022-4S-13-20
12. Shukhaev SV, Mordovtseva EA, Pustozerov EA, Kudlakhmedov SS. Application of convolutional neural networks to define Fuchs endothelial dystrophy. *Fyodorov Journal of Ophthalmic Surgery*. 2022;(4S):70–76. EDN: WEZTKV doi: 10.25276/0235-4160-2022-4S-70-76

ОБ АВТОРАХ

* **Арзамасцев Александр Анатольевич**, д-р техн. наук, профессор;
адрес: Россия, 392000, г. Тамбов, ул. Лаврова, д. 24;
ORCID: 0000-0001-6795-2370;
eLibrary SPIN: 4410-6340;
e-mail: arz_sci@mail.ru

Фабрикантов Олег Львович, д-р мед. наук, профессор;
ORCID: 0000-0003-0097-991X;
eLibrary SPIN: 9675-9696;
e-mail: fabr-mntk@yandex.ru

Кулагина Елена Валерьевна;
ORCID: 0009-0006-0026-0832;
eLibrary SPIN: 8785-4949;
e-mail: irina-kulagin2015@yandex.ru

Зенкова Наталья Александровна, канд. психол. наук, доцент;
ORCID: 0000-0002-2325-1924;
eLibrary SPIN: 2266-4168;
e-mail: natulin@mail.ru

AUTHORS' INFO

* **Alexander A. Arzamastsev**, Dr. Sci. (Engineering), Professor;
address: 24 Lavrov street, 392000 Tambov, Russia;
ORCID: 0000-0001-6795-2370;
eLibrary SPIN: 4410-6340;
e-mail: arz_sci@mail.ru

Oleg L. Fabrikantov, MD, Dr. Sci. (Medicine), Professor;
ORCID: 0000-0003-0097-991X;
eLibrary SPIN: 9675-9696;
e-mail: fabr-mntk@yandex.ru

Elena V. Kulagina, MD;
ORCID: 0009-0006-0026-0832;
eLibrary SPIN: 8785-4949;
e-mail: irina-kulagin2015@yandex.ru

Natalia A. Zenkova, Cand. Sci. (Psychology), Assistant Professor;
ORCID: 0000-0002-2325-1924;
eLibrary SPIN: 2266-4168;
e-mail: natulin@mail.ru

* Автор, ответственный за переписку / Corresponding author