

УДК 004.932:617.7

# **ГЛУБОКОЕ ОБУЧЕНИЕ ДЛЯ СКРИНИНГА ДИАБЕТИЧЕСКОЙ РЕТИНОПАТИИ: ОБЗОР ПОДХОДОВ ВАЛИДАЦИИ И МЕТРИК ПО ДАНЫМ КЛИНИЧЕСКИХ ПУБЛИКАЦИЙ**

**Петрова М.С.**

Независимый исследователь, Санкт-Петербург, Россия

E-mail: ms.petrova@example.com

## **Аннотация**

Цель работы - систематизировать ключевые элементы валидации алгоритмов глубокого обучения для обнаружения диабетической ретинопатии (ДР) на фотографиях глазного дна: качество данных, формирование обучающих и тестовых выборок, выбор порога решения, сравнение с референсной разметкой и представление диагностических метрик. В качестве основы рассмотрены крупные клинические публикации по разработке и проверке моделей, включая исследования Gulshan и соавт. (2016) и Ting и соавт. (2017), а также клиническое испытание автономной системы в первичном звене (Abramoff и соавт., 2018). Показано, что воспроизводимость результатов критически зависит от описания протокола разметки, контроля качества снимков и заранее заданной схемы оценки чувствительности и специфичности.

## **Ключевые слова**

диабетическая ретинопатия, глубокое обучение, изображения глазного дна, скрининг, чувствительность, специфичность, автономная диагностика

## **Введение**

Диабетическая ретинопатия остается одной из ведущих причин предотвратимой потери зрения, при этом эффективность профилактики зависит от своевременного выявления. Рост доступности цифровой фундус-фотографии и

развитие сверточных нейронных сетей привели к появлению алгоритмов, сопоставимых с экспертной оценкой по ключевым метрикам [1, 2]. Для внедрения таких систем в скрининг недостаточно высокой точности на одной базе данных - требуется прозрачная валидация на независимых выборках и корректное описание сравнения с референсным стандартом [3].

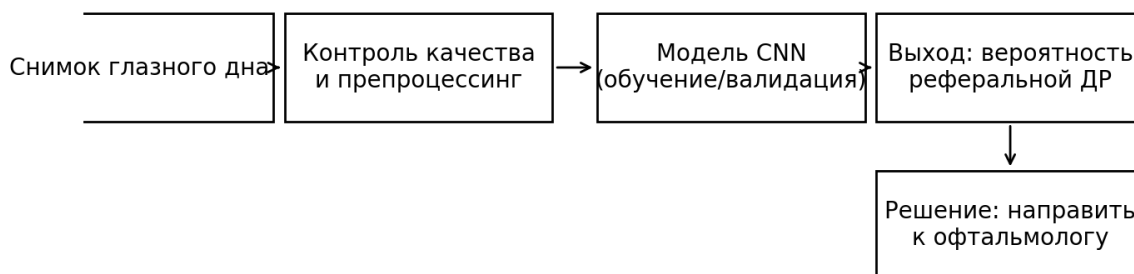
## Материалы и методы

В обзор включены клинические исследования и статьи по валидации алгоритмов обнаружения ДР на фотографиях глазного дна. При анализе фиксировались: источник и размер датасета, процедура градации ДР, критерии исключения изображений по качеству, а также метрики (AUC, чувствительность, специфичность). Особое внимание уделено публикациям в ведущих медицинских журналах и профильных изданиях по цифровой медицине [1-3].

**Таблица 1 - Примеры клинических исследований по валидации алгоритмов ДР**

Источник	Данные	Задача	Референс	Ключевые метрики
Gulshan и соавт., 2016 [1]	фундус-фото, большие датасеты	реферальная ДР	многоэкспертная разметка	чувствительность и специфичность на валидационных наборах
Ting и соавт., 2017 [2]	многоэтнические выборки	ДР и сопутствующие заболевания	клиническая оценка и грейдинг	оценка на данных разного качества
Abramoff и соавт., 2018 [3]	первичное звено, проспективно	автономная диагностика ДР	FPRC грейдинг + OCT	пороговые решения, клиническая применимость

**Рисунок 1 - Схема типового конвейера скрининга ДР на основе модели CNN**



## Результаты и обсуждение

Ключевой методологический узел - определение референсного стандарта. В клинических работах используется многоэтапная экспертная разметка, а при расхождениях - арбитраж, что снижает риск систематической ошибки в обучении и оценке [1]. Вторым важным фактором - контроль качества изображений: в реальных программах скрининга доля снимков, непригодных для оценки, выше, чем в тщательно отобранных датасетах, и это влияет на переносимость метрик [2]. В проспективных исследованиях автономных систем дополнительно важны операционные показатели: доля завершённых обследований, время получения результата и порядок направления пациента [3].

Для представления диагностической точности применяются базовые метрики классификации. Для бинарной задачи реферальной ДР удобно явно показывать формулы:

Чувствительность =  $TP / (TP + FN)$ ; Специфичность =  $TN / (TN + FP)$ .

При выборе порога решения необходимо указывать, оптимизировался ли он на обучающей выборке или задавался заранее. Отдельный вопрос - калибровка вероятностей: два алгоритма могут иметь одинаковую AUC, но разные свойства на клинически значимых порогах. В публикациях также важно указывать интервал доверия для метрик и схему бутстрэп-оценки, если она использовалась.

## **Заключение**

Исследования последних лет показывают, что глубокое обучение способно обеспечить высокую диагностическую точность при скрининге диабетической ретинопатии, а автономные решения могут быть применимы в первичном звене при соблюдении требований к данным и валидации. Для воспроизводимости и сопоставимости результатов необходимы стандартизированные протоколы контроля качества изображений, прозрачное описание референсной разметки и заранее заданные планы оценки метрик.

## **Конфликт интересов**

Автор заявляет об отсутствии конфликта интересов.

## **Финансирование**

Исследование выполнено без внешнего финансирования.

## **Список литературы**

1. Gulshan V., et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs // JAMA. 2016. Vol. 316(22). P. 2402-2410. DOI: 10.1001/jama.2016.17216.
2. Ting D.S.W., et al. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes // JAMA. 2017. Vol. 318(22). P. 2211-2223. DOI: 10.1001/jama.2017.18152.
3. Abramoff M.D., et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices // npj Digital Medicine. 2018. Vol. 1. Article 39. DOI: доступен на сайте Nature.
4. Keane P.A., Topol E.J. With an eye to AI and autonomous diagnosis // npj Digital Medicine. 2018. URL: <https://www.nature.com/articles/s41746-018-0048-y> (дата обращения: 03.03.2026).
5. Добров Э.Р., и др. Применение нейросетей в диагностике диабетической ретинопатии // Сборник материалов. 2022. URL: <https://mss.pnzgu.ru/files/mss.pnzgu.ru/07222.pdf> (дата обращения: 03.03.2026).

## **Дата поступления статьи в редакцию**

03.03.2026

**Дата принятия статьи к печати**

03.03.2026