

УДК 004.891

## ГИБРИДНЫЙ МЕТОД АНАЛИЗА ИЗОБРАЖЕНИЙ НА ОСНОВЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И НЕЧЕТКИХ МНОЖЕСТВ<sup>1</sup>

© 2025 г. А. Н. Аверкин<sup>а, \*</sup>, Е. Н. Волков<sup>а</sup>, С. А. Ярушев<sup>а</sup>

<sup>а</sup>Российский экономический университет им. Г.В. Плеханова, Москва, Россия

\*e-mail: averkin2003@inbox.ru

Поступила в редакцию 02.05.2024 г.

После доработки 08.11.2024 г.

Принята к публикации 24.02.2025 г.

Рассматривается разработка прототипа гибридной интеллектуальной системы для анализа изображений на примере задачи диагностики и стадирования диабетической ретинопатии — осложнения сахарного диабета, характеризующееся повреждением сосудов сетчатки глаза. В результате хронически повышенного уровня глюкозы в крови происходит нарушение микроциркуляции, что вызывает развитие микроаневризм, экссудации, кровоизлияний и в тяжелых случаях неоваскуляризации. Это может привести к ухудшению зрения и в конечном итоге к слепоте при отсутствии своевременного лечения. Выявление и стадирование заболевания происходит с помощью анализа фотографических изображений глазного дна (фундус-снимков). Проведен обзор по теме исследования, рассмотрены преимущества гибридных интеллектуальных систем в сравнении с решениями, основанными на применении одной технологии. Описаны шаги создания системы, сочетающей в себе совместное использование классических методов компьютерного зрения, искусственных нейронных сетей, элементов теории нечеткой логики и методов объяснительного искусственного интеллекта. С помощью комбинированной архитектуры программного решения удалось добиться гибкости в вопросах применимости критериев стадирования заболевания, что говорит о широких перспективах подобного решения в диагностике иных заболеваний с логически формализуемыми критериями.

*Ключевые слова:* гибридный метод, анализ изображений, искусственный интеллект, нечеткие множества, машинное обучение, компьютерное зрение, обработка изображений, нейронные сети, интеллектуальные системы, распознавание образов

DOI: 10.31857/S0002338825030103, EDN: BGWQJU

## HYBRID METHOD OF IMAGE ANALYSIS BASED ON ARTIFICIAL INTELLIGENCE TECHNOLOGIES AND FUZZY SETS

A. N. Averkin<sup>а, \*</sup>, E. N. Volkov<sup>а</sup>, S. A. Yarushev<sup>а</sup>

<sup>а</sup>Plekhanov Russian University of Economics, Moscow, Russia

\*e-mail: averkin2003@inbox.ru

The paper deals with the development of a prototype of a hybrid intelligent system for image analysis on the example of the task of diagnosis and staging of diabetic retinopathy — a complication of diabetes mellitus, characterized by damage to the retinal vessels. As a result of chronically elevated blood glucose levels, microcirculation is impaired, leading to the development of microaneurysms, exudation, hemorrhage and, in severe cases, neovascularization. This can lead to visual impairment and, ultimately, to blindness in the absence of timely treatment. Detection and staging of the disease are based on the analysis of photographic images of the ocular fundus (fundus images). An overview of the research topic is given, the basis for the advantages of hybrid intelligent systems in comparison with solutions based on the application of a single technology is presented. The steps of creating a system that combines the joint use of classical methods of computer vision, artificial neural networks, elements of fuzzy logic theory and methods of explainable

<sup>1</sup> Исследование выполнено за счет гранта Российского научного фонда (проект № 22-71-10112), <https://rscf.ru/project/22-71-10112/>

artificial intelligence are described. With the help of combined architecture of the software solution it was possible to achieve flexibility in the issues of applicability of criteria of disease staging, which indicates the broad prospects of such a solution in the diagnosis of other diseases with logically formalizable criteria.

*Keywords:* hybrid method, image analysis, artificial intelligence, fuzzy sets, machine learning, computer vision, image processing, neural networks, intelligent systems, pattern recognition

**Введение.** Повсеместная интеграция технологий искусственного интеллекта (ИИ) значительно улучшает результаты автоматизации подавляющего большинства процессов, в том числе и в медицине. Использование искусственных нейронных сетей (ИНС) в диагностике заболеваний на основе анализа медицинских изображений имеет широкую практику применения [1]. Однако для отдельных доменов медицинских изображений такая практика ограничена, в том числе для анализа офтальмологических изображений (рис. 1).

Прогресс в сфере обработки медицинских изображений играет ключевую роль в современной диагностике и исследованиях, обеспечивая возможность более точного определения диагнозов, выявления заболеваний на ранней стадии и планирования эффективного лечения. Среди многообразия подходов и методологий, применяемых в данной области, особое внимание привлекают гибридные методы анализа, которые интегрируют в себя достижения ИИ и теории множеств. Это сочетание предлагает обещающий потенциал для повышения точности и надежности анализа медицинских изображений за счет объединения вычислительной мощи ИИ для распознавания образов и адаптивности нечеткой логики для учета неопределенности, неточности, часто встречающихся в медицинских данных.

Несмотря на значительное количество исследований, посвященных отдельно прикладным аспектам ИИ в анализе изображений и применению нечетких множеств в медицине, лишь ограниченное количество работ рассматривает интеграцию этих двух направлений [2]. Гибридный подход, объединяющий ИИ и нечеткие множества, предоставляет уникальные возможности для обработки и интерпретации медицинских изображений, позволяя улучшить качество диагностики за счет более глубокого анализа и визуализации данных, а дополнение таких систем методом объяснения результата работы ИНС выводит их на совершенно новый уровень [3].

**1. Гибридные методы анализа медицинских изображений.** 1.1. Постановка задачи. Учитывая выбор гибридной архитектуры программного решения, целесообразно провести формальную постановку задачи, решаемой каждым из компонентов системы (рис. 2).

Определим, что поступающим на вход системы фундус-снимком будем считать цифровое изображение  $I$ , представляемое в виде матрицы целых чисел размера  $W \times H \times C$ , где  $W$  – ширина,  $H$  – высота,  $C$  – количество каналов. Примем, что для цветного изображения в формате RGB количество таких каналов  $C = 3$  с диапазоном значений цвета каждого пикселя  $i$  для каждого канала от 0 до 255 (рис. 3, 4).

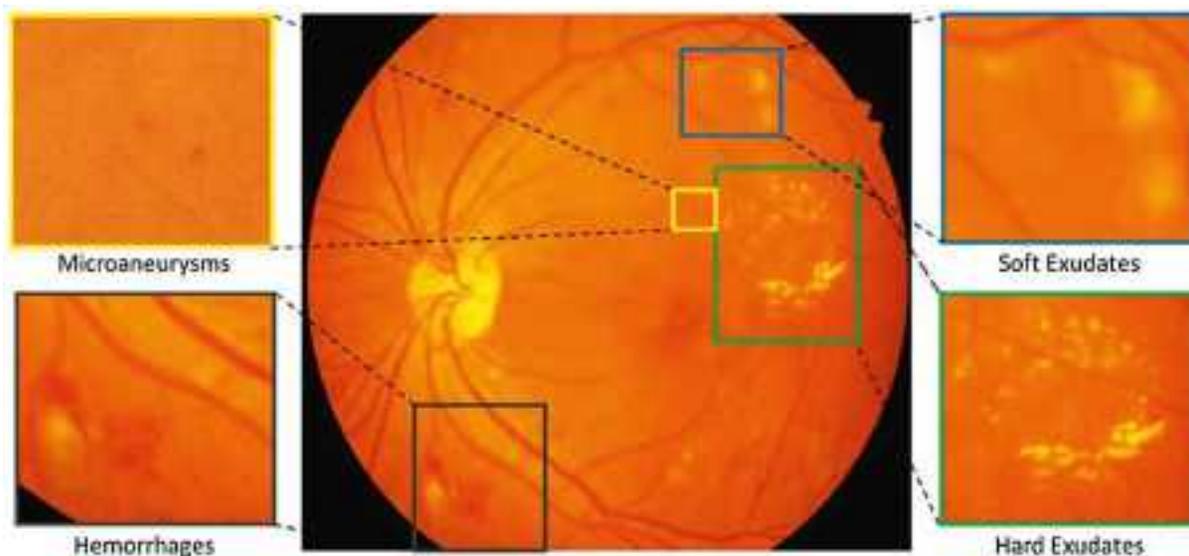


Рис. 1. Признаки диабетической ретинопатии на снимке глазного дна.

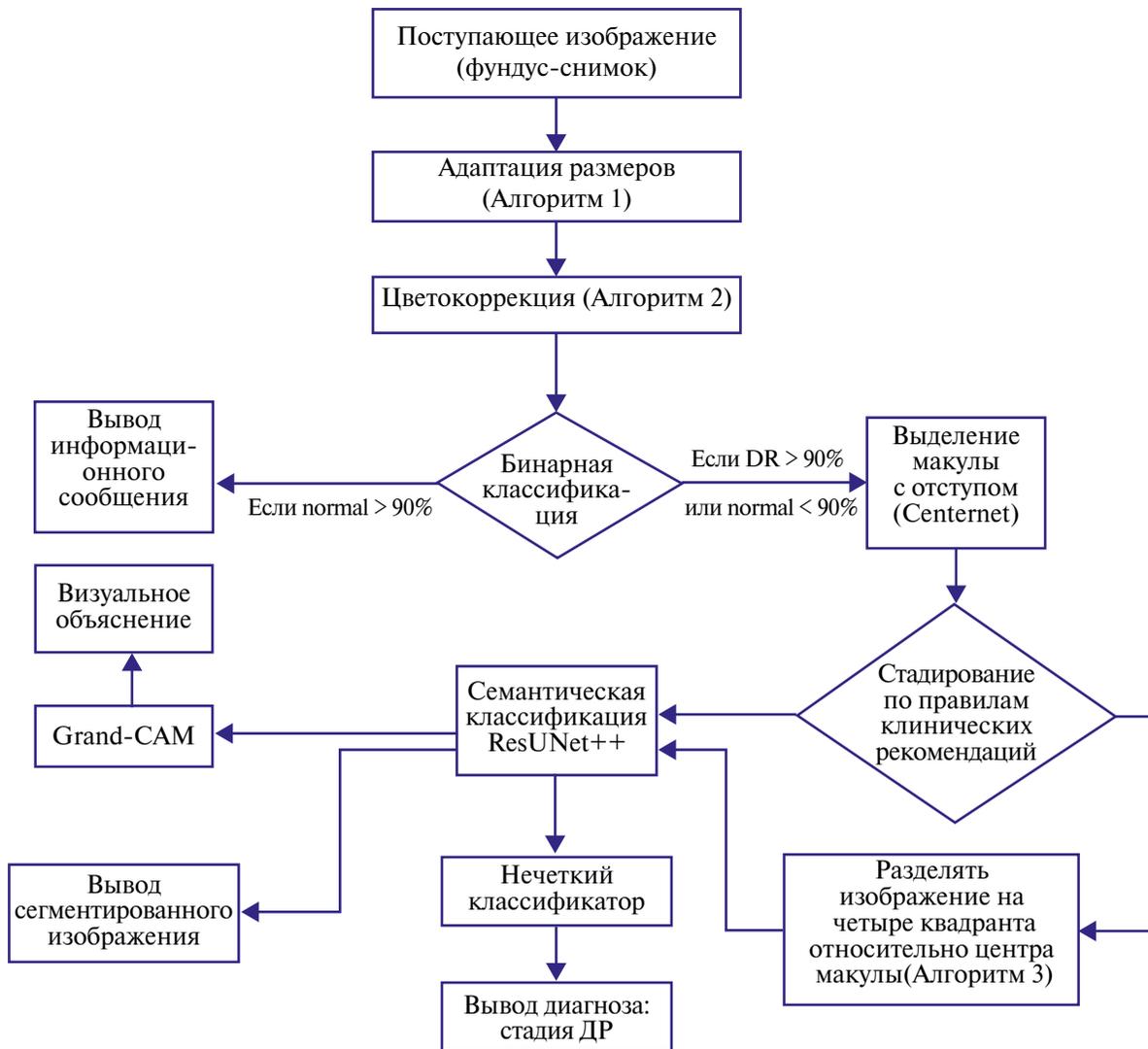


Рис. 2. Структура гибридной интеллектуальной системы диагностики диабетической ретинопатии.

1.1.1. Предподготовка изображений. Цель – коррекция входящего изображения, необходимая для корректной работы нейросетевых модулей и получения максимально точного предсказания, т.е. обобщения максимально возможного количества признаков с помощью ИНС. Предподготовка фундус-снимков имеет устоявшиеся стандарты, описанные во многих работах. Помимо стандартных шагов, применяемых во всех пайплайнах нейросетевого анализа изображения, таких как приведение размера входящего изображения к размеру входного слоя ИНС, нормализация значений пикселей в единый диапазон от 0 до 1, имеются несколько специфических, относящихся к конкретному домену изображения (рис. 3, 4).

1.1.2. Первичная классификация. Цель – отсеивать изображения без признаков искомого заболевания. Данный этап описан задачей бинарной классификации. Пусть дана выборка  $D = \{\vec{x}_i, y_i\}_{i=1}^N$ , где  $\vec{x}_i \in X = \mathbb{R}^m$  – множество объектов (изображений), представленных в векторном виде, а  $y_i \in Y = \{1, \dots, C\}$  – конечное множество меток классов,  $C$  – количество классов в выборке,  $N$  – размер выборки. Если  $y_i \in \{-1, +1\}$ , то такая классификация называется бинарной. Необходимо построить такую модель решающей функции  $a : X \rightarrow Y$ , приближающую  $u$  на всем подмножестве  $X$ . В данном случае в виде решающей функции выступит ИНС.

После получения значений метрик классификации для поданного на сеть изображения  $I_{prep}$  принимается решение о дальнейших шагах. Если объект классифицирован как “диабетическая ретинопатия” с вероятностью более 90% или “отсутствие заболевания” с вероятностью менее 90%, то он передается на следующий этап. Если объект отнесен к классу “отсутствие

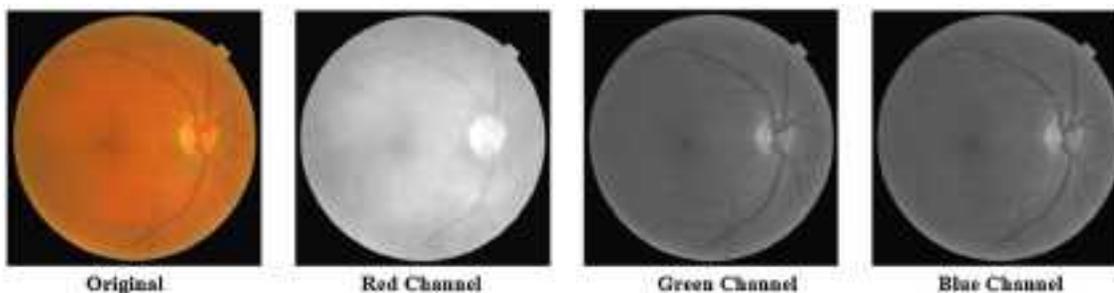


Рис. 3. Фундус-изображения в цветовых каналах RGB-диапазона (собственные данные).

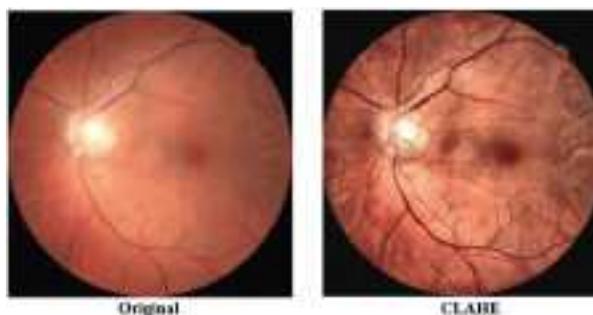


Рис. 4. Результаты применения фильтра CLAHE (собственные данные).

заболевания” с вероятностью более 90%, то анализ данного изображения прекращается и выводится информационное сообщение.

1.1.3. Сегментация патологий. Цель – выделить контуры и классифицировать участки, содержащие патологии (микроаневризмы, гемморагии, твердые и мягкие экструдаты). На данном этапе описана задача семантической сегментации.

Пусть каждому пикселю изображения  $I$ , представленного в виде матрицы размерами  $WH$ , соответствует некоторая дискретная функция  $u : \{1, M\} \times \{1, N\} \rightarrow \{0, 255\}^C$ , отражающая информацию об интенсивности  $i, j$  пикселей, где количество каналов  $C = 3$  для цветного изображения. Каждый пиксель отождествляется с соответствующими векторами  $x_j$ . Тогда сегментация изображения – это разбиение множества  $X = \{x_j\}_{j=1}^d$ , где  $d = WH$ , на кластеры  $X_j, j = \overline{1, l}$ , при котором каждому кластеру  $X_j$  присваивается определенное цветовое значение, соответствующее выделяемому им анатомической области изображения или фону.

По условию необходимо провести семантическую сегментацию, т.е. полностью разделить множество всех пикселей изображения  $P_I$  на  $n$  кластеров пикселей и определить принадлежность каждого из кластеров к какому-либо классу (1.1):

$$\bigcup_{i=1}^c X_j = P_I X_i \cap X_j \neq \emptyset, i \neq j, i, j = \overline{1, C}. \quad (1.1)$$

При этом в качестве решающей функции выступает ИНС. После получения результатов сегментации (обнаруженных классов и значений их вероятностей при переходе значений порога в 80%) они передаются на следующий этап.

1.1.4. Нечеткая классификация на основе правил. Цель – классификация патологии на основе формализованных в правила критериев постановки диагноза. Качество выполнения нечеткой классификации объектов зависит как от качества их описания, так и от качества формализации правил. Используя подход, введенный в работе [4], рассмотрим постановку задачи нечеткой классификации на базе правил в диагностике заболеваний с помощью анализа изображений. Пусть классифицируемый объект (изображение)  $A$  представляется неким конечным набором признаков  $A = \{A_1, \dots, A_n\}$ , извлеченных из него с помощью ИНС. Каждому признаку  $A_i$  ставится в соответствие множество  $U_i$  характеристик его описания (в данном случае количественных значений) и множество  $\{a_{i1}, \dots, a_{in}\}$  ( $1, n$ ) лингвистических значений. Каждому такому лингвистическому значению  $a_{ij}$  ставится в соответствие функция

принадлежности  $\mu_{a_{ij}}(u_i)$  на универсальном множестве  $U_i(\overline{1, n})$ . Кроме этого, имеется  $K$  классов  $C_k, k \in \{1, \overline{K}\}$ , в данном исследовании представленных стадиями диабетической ретинопатии. Информация о классах рассмотрена в виде совокупности правил  $r$ , полученных из формализованных критериев диагностики заболевания (клинические рекомендации, публикации и т.д.).

1.2. Описание наборов данных. Решение задач глубокого обучения невозможно без правильного выбора и обработки наборов данных для обучения ИНС. Подбор наборов данных в задачах анализа медицинских изображений представляет собой трудоемкий процесс, поскольку специфика домена изображений подразумевает, что для их разметки необходимо обладать исключительными знаниями и богатой экспертизой в диагностике конкретного заболевания. Поиск наборов данных фундус-изображений под разные типы задач (классификация, сегментация) также осложнен малым количеством наборов данных, находящихся в открытом доступе. Так, в результате исследования доступности наборов данных фундус-изображений [5] на 2023 г. всего представлено 120 наборов изображений, из которых 43 набора по тематике заболевания “диабетическая ретинопатия”, но только около трети наборов находятся в открытом доступе.

В качестве набора данных для бинарной классификации был выбран APTOS2019, состоящий из 3662 изображений всех четырех стадий диабетической ретинопатии (Mild DR – 370, Moderate DR – 999, Severe DR – 193, Proliferative DR – 295), а также изображений без признаков заболевания (No DR – 1805). Примеры изображений набора данных APTOS2019 представлены на рис. 5.

В оригинальном состоянии набор данных является несбалансированным, что может оказать негативное влияние на результаты классификации. Набор данных был сбалансирован по классам путем добавления примеров (аугментации) с помощью поворота изображения на угол от 45 до 270° в произвольном порядке, поскольку данный вариант изменения изображения не несет в себе опасности потери или искажения важных признаков. Учитывая постановку задачи, набор данных был преобразован в набор для бинарной классификации путем объединения всех изображений, относящихся к какой-либо из стадий заболевания в один класс. В дальнейшем к получившемуся новому набору данных был применен метод предподготовки перед подачей на обучение ИНС. Характеристики набора данных для бинарной классификации представлены в табл. 1.

По сравнению с наборами данных для задачи классификации диабетической ретинопатии, наборы данных для решения задачи сегментации патологий, на основе наличия которых устанавливается диагноз “диабетическая ретинопатия” (микроаневризмы, кровоизлияния,

Таблица 1. Набор данных для классификации

Классы		Стадии	Соотношение изображений (количество/процент)				
Бинарный	Мультиклассовый		APTOS2019 (оригинальный)		Аугментация	Балансировка	
			Кол-во	%		Кол-во	Кол-во
<i>Normal</i>	<i>Normal</i>	–	1805	49.3	+595	2400	50
<i>Diabetic Retinopathy</i>	<i>Mild</i>	1	370	10.1	+230	2400	50
	<i>Moderate</i>	2	999	27.28	–399		
	<i>Severe</i>	3	193	5.27	+407		
	<i>Proliferative</i>	4	295	8.05	+305		
<b><i>Всего</i></b>			<b>3662</b>	–	<b>1537</b>	<b>4800</b>	–



Рис. 5. Примеры изображений набора данных APTOS2019.

твердые и мягкие экссудаты), количество общедоступных наборов данных ограничено. В работе были использованы два наиболее известных набора данных для сегментации фундус-изображений FGADR и IDRID (рис. 6).

Набор данных FGADR состоит из 2842 изображений с разрешением 1280 на 1280 пикселей, разделенных на два поддета: для классификации стадии – 1000 изображений и сегментации патологий (microaneurysms, hemorrhages, hard exudates, soft exudates, intra-retinal microvascular abnormalities, neovascularization) – 1842 изображений и масок классов. Другой набор данных – IDRID состоит из 81 изображения патологий с разрешением 4288 на 2848 и масок классов к ним (hard exudates, soft exudates, microaneurysms, hemorrhages).

1.3. Выбор архитектуры ИНС. Принимая во внимание необходимость решения трех совершенно разных задач анализа изображения (классификация, детекция, сегментация), необходимо выбрать оптимальные по соотношению затрат вычислительных мощностей и качества предсказаний архитектуры ИНС. Учитывая условия, лучших результатов можно добиться при использовании готовых архитектур, предобученных на больших массивах данных, с дальнейшим применением техники трансферного обучения для адаптации моделей под задачи исследования.

1.3.1. Бинарная классификация. EfficientNetV0 – это модель глубокого обучения, разработанная для задач классификации изображений, которая относится к семейству архитектур EfficientNet. Эти модели были представлены в 2019 г. в исследовании, проведенном Google Research, и быстро завоевали популярность за свою высокую эффективность и относительно небольшой размер (рис. 7).

Один из ключевых принципов, лежащих в основе EfficientNet, – вместо того, чтобы просто увеличивать глубину (количество слоев) или ширину (количество узлов в слое), или разрешение входного изображения, эффективное масштабирование требует одновременного и сбалансированного увеличения всех трех измерений, используя составной коэффициент масштабирования. Для EfficientNetV0 данный подход был применен впервые, а последующие версии (от V1 до V7) использовали большие значения этого коэффициента для достижения лучшей производительности.

В основе EfficientNet лежат блоки MBConv – это улучшенная версия инвертированных остаточных сверточных блоков, которые впервые были представлены в MobileNetV2. Эти блоки используют механизм squeeze-and-excitation для повышения эффективности обработки сигналов внутри блока, а также применяют линейные остаточные соединения и инвертированные остатки для улучшения потока градиентов в глубоких сетях. Свертки с определением шага (strided convolutions) и отдельные свертки (depthwise convolutions) используются для снижения вычислительных требований и уменьшения объема параметров в сети, сохраняя при этом качество обработки изображений.

EfficientNetV0 достигает высокой точности при значительно меньшем количестве параметров и вычислительной мощности, чем другие архитектуры сетей того же времени, что делает ее одним из предпочтительных выборов для задач классификации изображений, особенно

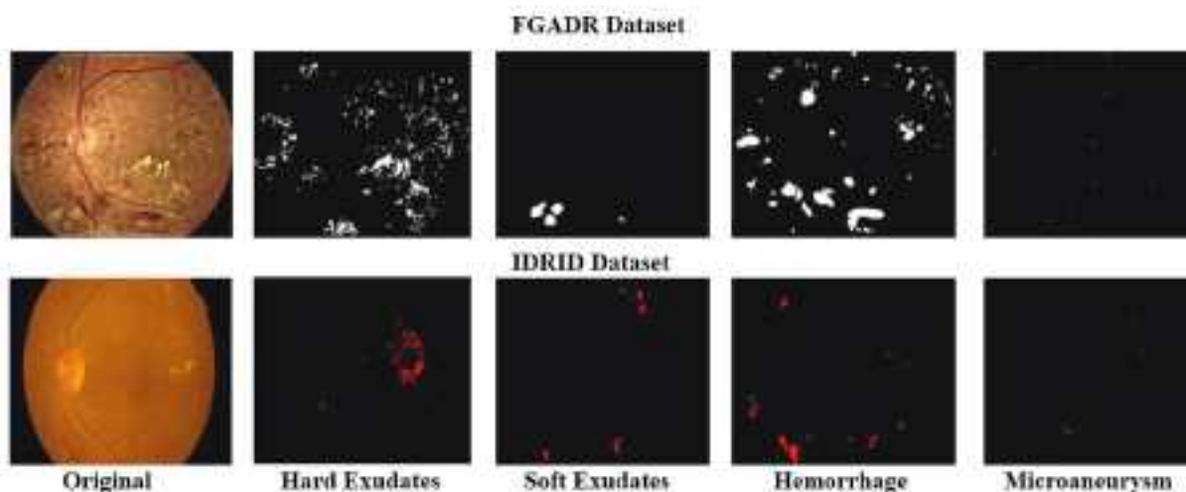


Рис. 6. Примеры изображений наборов данных для сегментации: первая строка – FGADR, вторая строка – IDRID, столбцы – маски классов.

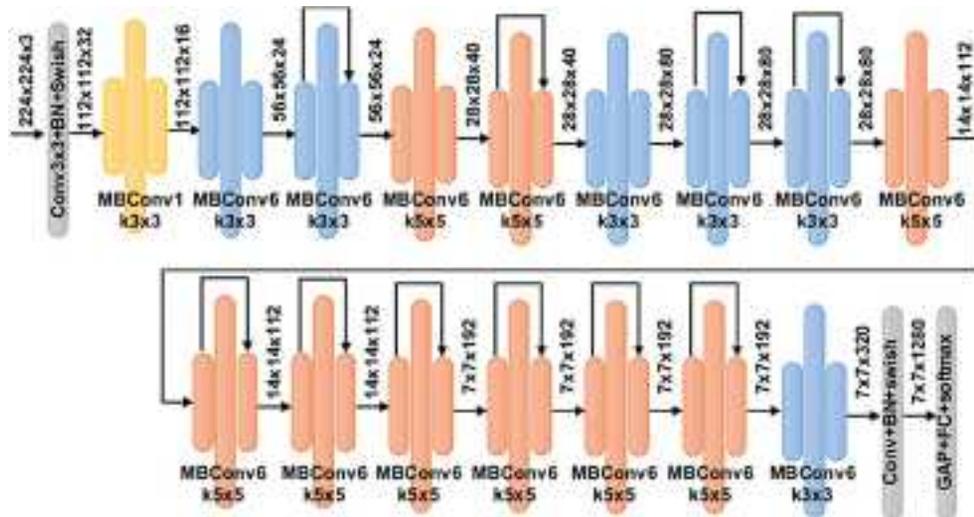


Рис. 7. Архитектура EfficientNetB0.

в условиях ограничения ресурсов. Сети этого семейства нашли широкое применение в анализе медицинских изображений. Дополнительно для дальнейшего сравнения результатов в качестве базовой модели была выбрана архитектура VGG-16.

1.3.2. Детекция. Архитектура CenterNet является методом обнаружения объектов, который в отличие от традиционных методов, основанных на якорных (anchor-based) или региональных (region proposal-based) подходах, сосредотачивает внимание на обнаружении центральных точек объектов и использует тройки ключевых точек для определения положения, размера и класса объектов.

CenterNet задает каждый объект как тройку ключевых точек: центр, верхний правый угол и нижний левый угол прямоугольника, описывающего объект. Эта схема позволяет точно найти положение и размер объекта без применения предварительно заданных якорей. Для определения размера объектов CenterNet использует предсказания размеров ограничивающего прямоугольника относительно найденного центра объекта, что позволяет точно получать размеры объектов, независимо от их положения на изображении. В свою очередь, для обнаружения центральных точек объектов CenterNet генерирует карту активации, где каждому типу объекта соответствует свой слой. Значение в каждой точке этой многомерной карты отражает вероятность того, что точка является центром объекта данного класса. CenterNet представляет собой продвинутый и эффективный подход к задаче обнаружения объектов, который обеспечивает высокую точность при упрощенной архитектуре (рис. 8).

1.3.3. Семантическая сегментация. В настоящее время существует множество архитектур ИНС, созданных для решения задач сегментации изображений. Однако наиболее популярными среди них являются нейросети, основанные на архитектуре типа U-Net. В данной работе будет использоваться одна из модификаций классической сверточной

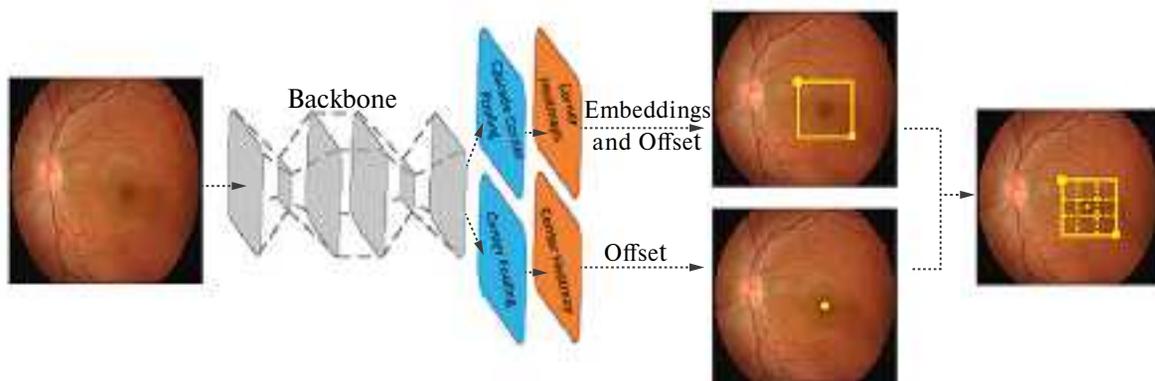


Рис. 8. Архитектура CenterNet.

архитектуры – ResUNet++, созданная специально для решения задач сегментации медицинских изображений. В качестве базовой модели для решения данной задачи была выбрана классическая U-Net.

Архитектура ResUNet++ была предложена в 2019 г. и представляет собой усовершенствованную версию архитектуры ResUNet, которая в свою очередь является гибридом между ResNet и U-Net для задач сегментации изображений преимущественно в медицинской области [6]. Эта архитектура используется для улучшения точности сегментации за счет более сложного и эффективного извлечения признаков и реконструкции таргетного изображения на выходе нейросети, что стало возможным благодаря применению резидуальных блоков и механизма внимания (рис. 9).

Блоки U-Net состоят из кодировщика и декодировщика с соединениями пропуска (skip connections). Структура напоминает букву “U”, где кодировщик последовательно уменьшает размерность карты признаков, а декодировщик ее восстанавливает. Соединения пропуска помогают восстановить детальную информацию о форме и расположении объектов на изображении, что критично для задач сегментации.

Сверточная база ResNet используется для первичного извлечения признаков. Модели ResNet (Residual Networks) примечательны своими остаточными соединениями (skip connections), которые помогают уменьшить проблему исчезающего градиента при обучении глубоких сетей, предоставляя прямой путь прохождения градиентов. Блоки механизма внимания добавлены для более эффективного фокусирования на важных признаках изображений (блоки Attention на рис. 9). Механизмы внимания могут быть реализованы различными способами, например через каналное внимание, пространственное или внимание, заточенное под конкретные задачи. Эти механизмы помогают модели лучше концентрироваться на ключевых деталях и игнорировать нерелевантный фон.

Дополнительно в сеть были интегрированы блоки сжатия и возбуждения (SE-блоки) из SENet, блоки агрессивного пространственного пирамидного пула (ASPP) из DeepLabv3. ASPP позволяет собирать информацию из входных данных на разных разрешениях благодаря использованию сверток с различными степенями разреженности (atrous convolutions). Взвешивание каналов в SE-блоках дает возможность модели находить важность каждого канала признаков. В результате сеть может автоматически определять и уделять больше внимания более информативным признакам, в то время как менее важные признаки подавляются. Благодаря возможности адаптации к различным типам признаков SE-блоки способствуют улучшению распознавания контекста изображения. Это особенно важно в медицинских задачах сегментации, где необходимо точно отличать ткани или анатомические структуры друг от друга. Это повышает способность сети к более точному выделению интересующих объектов на изображении.

ResUNet++ нашла широкое применение в медицинской сегментации, например в автоматической сегментации опухолей, органов или тканей на медицинских сканах. Модернизированная и переработанная версия классической архитектуры U-Net значительно улучшает показатели качества сегментации, что подтверждается как численными экспериментами авторов оригинальной статьи, так и большим количеством исследований.

1.4. Нечеткий классификатор. Согласно постановке задачи, установим, что набор совокупности  $r$  правил представлен в следующем виде:

$$\left\{ \begin{array}{l} \text{если } A_1 = a_{i_1}^1 \text{ и } A_2 = a_{i_2}^1 \text{ и } \dots \text{ и } A_n = a_{i_n}^1, \text{ то } \mu_{C_1}(O) = \eta_1^1 \text{ и } \dots \text{ и } \mu_{C_k}(O) = \eta_k^1, \\ \dots \\ \text{если } A_1 = a_{i_1}^r \text{ и } A_2 = a_{i_2}^r \text{ и } \dots \text{ и } A_n = a_{i_n}^r, \text{ то } \mu_{C_1}(O) = \eta_1^r \text{ и } \dots \text{ и } \mu_{C_k}(O) = \eta_k^r, \end{array} \right.$$

где объект  $O$  имеет значения признаков, перечисленных в левой части правил,  $\eta$  – степень уверенности того, что объект  $O$  принадлежит определенному классу в соответствии с  $r$  – м правилом,  $\mu$  – функция принадлежности, которая задает степень принадлежности объекта  $O$  к некоторому классу. Тогда назовем полученную систему (1.2) нечетким классификатором на основе совокупности  $r$  правил, составленных из формализованных критериев диагностики заболевания. На вход классификатора будет подаваться  $\mathbf{u}^* = \{u_1, \dots, u_n\}$ , где  $\mathbf{u}^* \in \mathbb{R}$  и является набором характеристик (значений) описания признаков объекта. Выходом служит  $\mathbf{M}^* = (M_1, \dots, M_K)$ , описывающий принадлежность объекта (изображения) классам  $C_1, C_K$ .

1.5. Формирование совокупности правил. Качество полученного вывода при использовании нечеткого классификатора напрямую зависит от качества формирования нечетких правил. По сути, формирование правил является процедурой формализации экспертного знания в форму, требующуюся классификатору для работы. Применительно к задаче

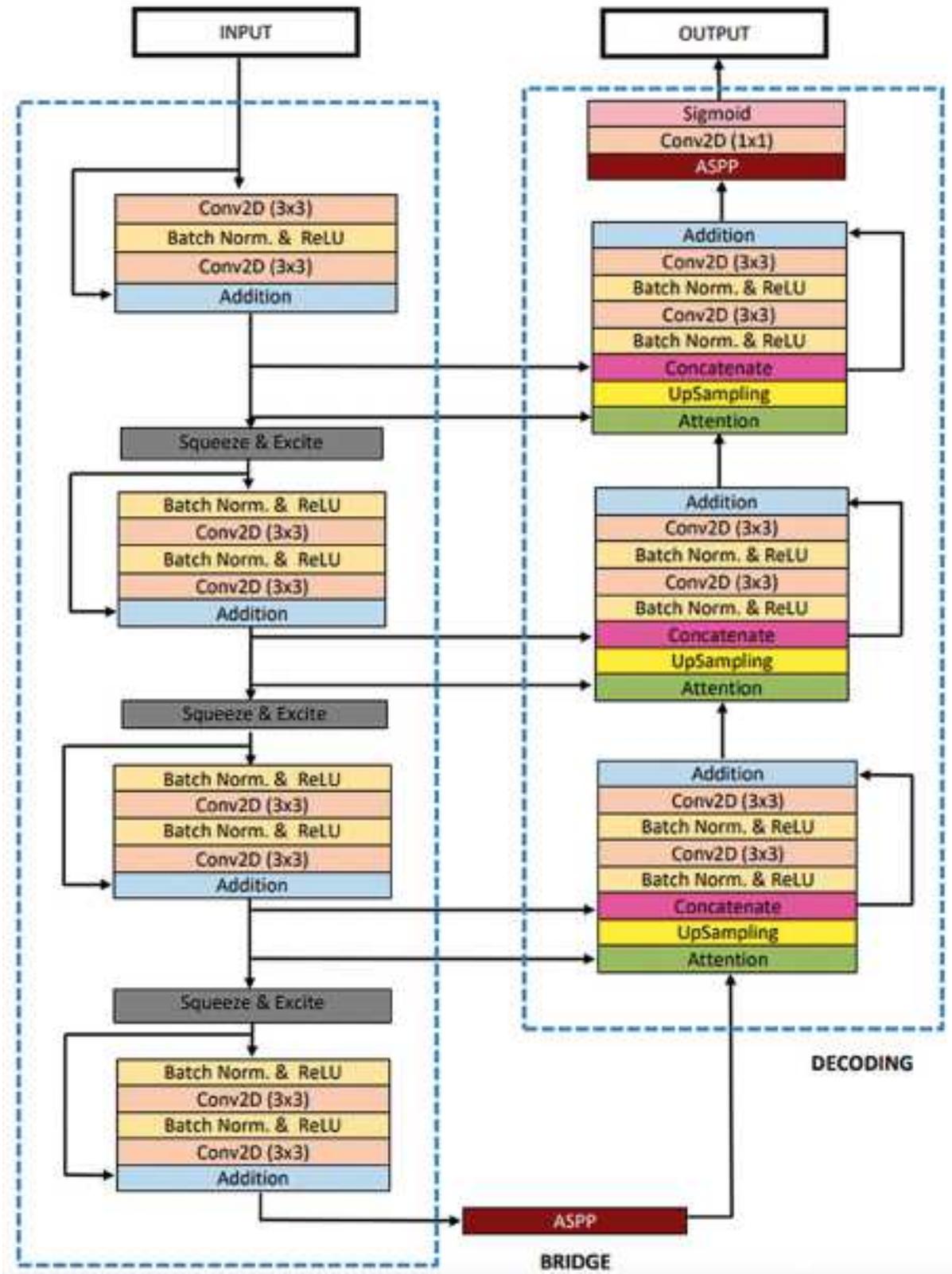


Рис. 9. Архитектура ResUNet++ [3].

диагностики заболеваний целесообразным и допустимым будет извлечение экспертных знаний из клинических рекомендаций, многоцентровых исследований. Однако, учитывая ограничения, связанные с доступностью данных, используемых для обучения ИНС, правила также могут быть извлечены из иных научных публикаций.

В Российской Федерации диагностика и стадирование диабетической ретинопатии по снимкам фундус-камеры производится на основе стандартов, установленных клиническими рекомендациями. Согласно приведенным клиническим рекомендациям, диабетическая ретинопатия подразделяется на три стадии: непролиферативная (НПР), препролиферативная (ПРПР), пролиферативная (ПР). Тяжесть заболевания возрастает от первой к третьей стадии. Все стадии диагностируются на основании выявления отдельных деструктивных изменений анатомических структур глаза, таких как микроаневризмы (МА), ретинальные кровоизлияния (НЕМ), твердые (НТ) и мягкие (СМ) экссудаты, разновидности неоваскуляризации (НВ), интратретинальных микрососудистых аномалий (ИРМА, ИРМА).

Представленные критерии соответствуют наполнению используемого для сегментации набора данных, однако требуют дополнительного разделения изображения на квадранты. В связи с этим для проведения экспериментов критерии были подвергнуты упрощению.

На основе представленных в табл. 2. характеристик изменений глазного дна был формализован набор из 135 нечетких правил, составленных для стадии НПР:

If (МА is слабые) and (НЕМ is слабые) and (СМ is низкая), then (НПР is низкая);

...

If (МА is слабые) and (НЕМ is средние) and (СМ is средняя), then (НПР is средняя);

...

If (МА is сильные) and (НЕМ is высокая) and (СМ is высокая), then (НПР is высокая).

Имитационное моделирование проводилось в среде MATLAB, а программная реализация – в виде модуля системы на языке программирования Python. Получившаяся поверхность принятия диагностического решения для НПР рассмотрена на рис. 10.

В свою очередь авторы набора данных Messidor изначально объявили критерии отнесения фундус-снимков к каждой из стадий диабетической ретинопатии. Для этого были введены следующие обозначения:  $\mu A$  – количество микроаневризм,  $H$  – количество ретинальных кровоизлияний,  $NV$  – наличие неоваскуляризации. Патологии, составляющие данные критерии, полностью совпадают с классами набора данных для сегментации изображений, разбиения изображения на квадранты не требуется. Найденные в результате формализации правил лингвистические критерии представлены в табл. 3, на основе которых были составлены 108 нечетких правил.

1.6. Выбор метода получения визуального объяснения. Определение визуального объяснения работы ИНС как априори непрозрачного алгоритма является одним из ключевых шагов в создании интерпретируемого программного решения. Использование методов объяснительного ИИ (ОИИ) в нейросетевом анализе медицинских изображений помогает увеличить уверенность конечного пользователя – врача в правильности предсказания за счет визуального выделения зон изображения, внесших наибольший вклад в предсказание.

Применение методов ОИИ для получения дополнительного вывода моделью визуального объяснения имеет широкую практику использования в анализе многих доменов медицинских изображений (компьютерная томография (КТ), магнитно-резонансная томография (МРТ), рентген-снимки, снимки кожи и т.д.) [7]. Наиболее популярно применение ОИИ в анализе снимков, полученных с помощью методов лучевой диагностики, что во многом объясняется как распространенностью томографов, так и способностью выгрузки их снимков в едином формате [8]. В свою очередь использование методов ОИИ до сих пор не нашло широкого распространения в анализе офтальмологических изображений (снимки фундус-камеры, снимки оптической компьютерной томографии (ОКТ)). Предыдущие исследования [3, 9] показывают,

**Таблица 2.** Формирование правил согласно клиническим рекомендациям РФ

Стадия	Характеристики изменений глазного дна
НПР	Микроаневризмы, ретинальные кровоизлияния, мягкие экссудаты
ПРПР	Наличие хотя бы одного из трех признаков: наличие ИРМА хотя бы в одном квадранте; венозные аномалии в более чем двух квадрантах; ретинальные кровоизлияния во всех четырех квадрантах
ПР	Неоваскуляризация (и/или фиброзная пролиферация) диска зрительного нерва и/или сетчатки, преретинальные и/или витреальные кровоизлияния, тракционная (или тракционно-регатогенная) глаукома сетчатки

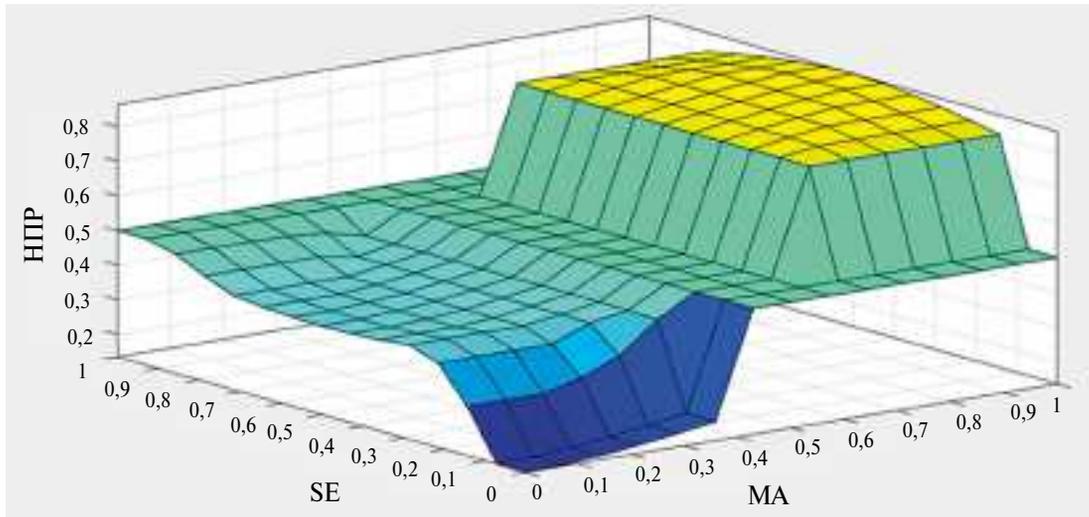


Рис. 10. Поверхность принятия диагностического решения для стадии диабетической ретинопатии – НПР (собственные данные).

Таблица 3. Критерии из набора данных Messidor

Стадия	Критерии стадирования диабетической ретинопатии
0	$(\mu A = 0) AND (H = 0)$
1	$(0 < \mu A \leq 5) AND (H = 0)$
2	$((0 < \mu A \leq 15) OR (0 < H < 5)) AND (NV = 0)$
3	$(\mu A \geq 15) OR (H \geq 5) OR (NV = 1)$

что методы ОИИ в основном применяются в задачах классификации фундус-снимков, в большей степени диабетической ретинопатии. Использование методов ОИИ в задачах сегментации заболеваний глаза носит крайне ограниченный характер.

В настоящее время, согласно метаобзорам [10, 11], описано более сотни методов ОИИ. Основные их различия составляют принципы получения объяснений и тип данных, для которых его возможно объяснить. В подавляющем большинстве случаев применяются методы, с помощью которых объясняются результаты работы непрозрачных моделей – ”post hoc” объяснения. В целом, методы ОИИ можно разделить на внешние интерпретаторы – “model agnostic” или встраиваемые в архитектуру сети – “model specific”. Наиболее предпочтительны в программных решениях – встраиваемые методы, поскольку они менее требовательны к вычислительным мощностям и затрачивают меньше времени для вывода объяснения.

Одним из наиболее часто используемых для получения визуального объяснения методов ОИИ является Grad-CAM (gradient-class activation mapping). Grad-CAM был предложен в 2017 г. [12] как обобщенная версия появившегося ранее метода CAM (class activation mapping) [13] с модификациями, благодаря которым не требуется внесения изменений в архитектуру сети, что достигается избавлением от линейности в финальной части интерпретируемой модели.

Наиболее простой метод визуализации объяснения работы ИНС – показать активацию интерпретируемой сети во время прямого прохода (процесса получения предсказания), а не во время обучения. Визуализации активации возможна через создание карт активации – визуального представления активированных нейронов на различных слоях сети. Пусть  $A^k$  – карты признаков, а  $y^c$  – выходные данные сети до применения Softmax. Вычислим градиент оценки искомого класса  $y^c$  относительно карты признаков  $A^k$ , затем усредним относительно размерности карт активации  $Z$  и определим веса важности нейрона  $w_k^c$ . На финальном этапе происходит применение функции активации RELU, позволяющей отфильтровать только положительные значения, т.е. только те значения, которые вносят какой-то вклад в итоговое

предсказание модели. Выводится визуальное объяснение в виде тепловой карты, показывающей интенсивностью цвета вклад отдельных регионов в предсказание.

Применение Grad-CAM в задаче семантической сегментации позволит получить визуальное объяснение сегментированных областей в виде слоя тепловой карты, наложенной поверх сегментированных изображений. Наиболее ярким цветом будут отмечены области, внесшие наибольший вклад в итоговое предсказание.

1.7. Метрики оценки качества. Выбор метрик оценки качества работы ИНС зависит, во-первых, от решаемой задачи, а во-вторых, от распределения данных в используемых наборах, поскольку дисбаланс классов может привести к некорректной работе модели на слабо представленных классах. Для подзадачи бинарной классификации, учитывая сбалансированность полученного набора данных, целесообразно применить метрики accuracy и F1-score. При этом ключевой метрикой в данной подзадаче будем считать accuracy.

Accuracy – доля объектов, для которых моделью правильно предсказан класс. Приведенная метрика строится на основе значений true positive (TP), false positive (FP), true negative (TN), false negative (FN), отражающих соотношение истинного и прогнозируемого класса объекта. Таким образом, характеризуем качество модели, агрегированное по всем классам:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (1.3)$$

F1-score – это гармоническое среднее, рассчитываемое на основе значений точности (Precision) и полноты (Recall). Точность – мера того, сколько объектов, определенных как положительные, реально ими являются. Полнота – мера того, сколько экземпляров конкретного класса были классифицированы как экземпляр этого класса:

$$Precision = \frac{TP}{TP + FP}, \quad (1.4)$$

$$Recall = \frac{TP}{FN + NP}, \quad (1.5)$$

$$F1 = \frac{2PrecisionRecall}{Precision + Recall}. \quad (1.6)$$

Отдельно выделяются метрики для задачи сегментации. В задачах сегментации оценка качества работы классификатора базируется на совпадении истинного (маска классов сегментации) и предсказанного на изображении региона пикселей. Для оценки качества сегментации используем две метрики: IoU и Dice. Однако будем считать коэффициент Дайса ключевой метрикой, поскольку он учитывает как истинно положительные, так и ложноположительные показатели, обеспечивая сбалансированную оценку эффективности сегментации [14].

Intersection over Union (IoU), также известный как индекс Жаккарда, – это площадь пересечения прогнозируемого региона изображения с истинным регионом, которая содержит искомый объект, деленная на общую площадь обоих регионов. Пороговым значением для исключения сомнительных прогнозов может служить порог отсечения 0.5 и 0.95:

$$IoU = \frac{area_{intersection}}{area_{union}}, \quad (1.7)$$

где  $area_{intersection}$  – площадь пересечения регионов, а  $area_{union}$  – площадь объединения регионов. Однако, согласно результатам отдельных исследований, данная метрика может быть излишне чувствительна к ошибкам границ для больших объектов и может чрезмерно наказывать за ошибки на меньших объектах [15].

Коэффициент сходства Дайса (DSC) изначально определяется как схожесть двух некоторых наборов данных, но в отношении задачи сегментации является схожестью истинных и предсказанных множеств значений пикселей изображения:

$$DSC = \frac{2TP}{(TP + FP) + (TP + FN)}. \quad (1.8)$$

Чем ближе значение коэффициента к единице, тем выше сходство сегментированного объекта с истинным. Коэффициент Дайса также имеет недостатки. Например, в исследовании [16] показаны проблемы использования данной метрики при наличии пересечений экземпляров одного класса.

**2. Результаты и обсуждение.** 2.1. Обучение и валидация ИНС. Обучение и валидация используемых в работе архитектур ИНС (EfficientNetB0, CenterNet, ResU-Net++) производилась на наборах данных, разделенных на обучающую и валидационную выборки в соотношении 80 на 20%. Контроль корректности работы сетей проходил на валидационных частях

выборок. Подготовка базовых моделей для сравнения результатов была произведена на тех же условиях. Результаты, представленные в табл. 4–6, свидетельствуют о превосходстве выбранных моделей ИНС по показателям над базовыми решениями.

Для улучшения результатов работы ResUNet++, помимо стандартных методов регуляризации, была применена аугментация во время тестирования (ТТА), которая является техникой, используемой для улучшения производительности моделей машинного обучения на этапе тестирования или вывода. Суть метода заключается в применении различных видов аугментации (увеличения) данных к тестовым образцам и последующем преобразовании выводов модели для получения окончательного результата. Например, изображение может быть масштабировано, повернуто или отражено перед его подачей в модель, и далее результаты для всех модификаций могут быть агрегированы (например, усреднены) для получения окончательного предсказания. Это может помочь улучшить точность модели, так как она демонстрирует более высокую обобщающую способность, учитывая разнообразные вариации входных данных. Результаты, найденные для тестовой выборки, представлены в табл. 6, а маски сегментированных классов – на рис. 11.

**Таблица 4.** Результаты бинарной классификации

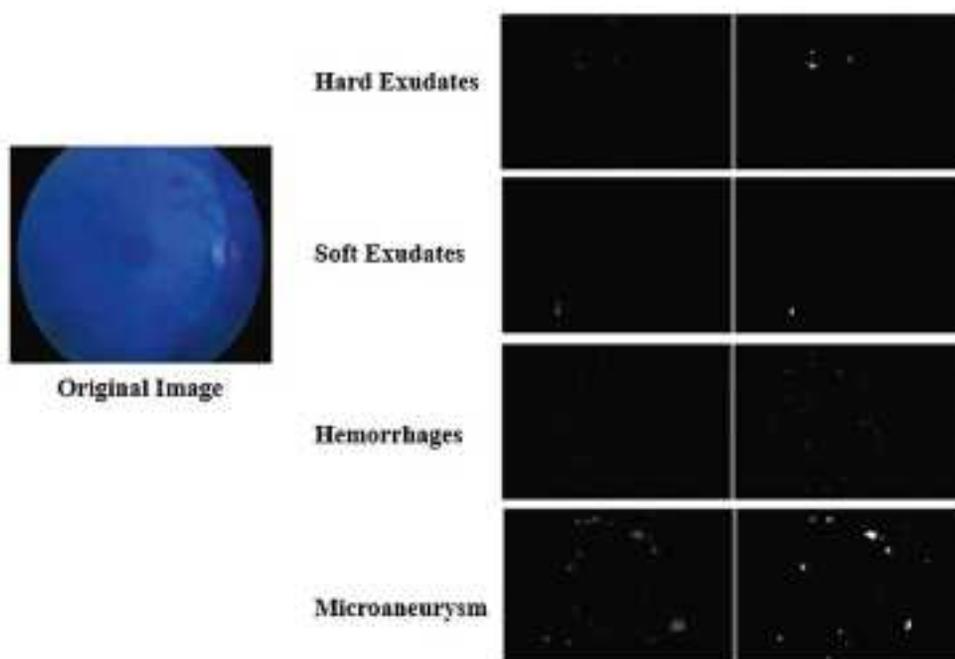
Модель	Accuracy, %	F1-score, %
Base (VGG-16)	0.93	0.91
EfficientNetB0	0.98	0.97

**Таблица 5.** Результаты детекции

Модель	Precision, %	IoU, %
Base (Yolov3)	0.84	0.82
CenterNet	0.93	0.9

**Таблица 6.** Результаты семантической сегментации

Модель	DSC, %	IoU, %
Base (U-Net)	0.67	0.71
ResUNet++	0.86	0.88
ResUNet++ и ТТА	0.89	0.9
ResUNet++ и CRF	0.91	0.87
ResUNet++ и ТТА и CRF	0.94	0.97



**Рис. 11.** Результаты сегментации с помощью ResUnet++ (собственные данные).

**Заключение.** Рассмотрен метод создания гибридных интеллектуальных систем на основе ИИ и нечеткой логики. Представлено программное решение на базе ИНС архитектур EfficientNetV0, CenterNet, ResNet++. Данное решение составляет комплексный метод диагностики диабетической ретинопатии по снимкам фундус-камеры и превосходит решения, состоящие только из одной технологии. Метод является универсальным и может быть применен во всех задачах, где можно получить описание предмета с использованием нечетких множеств.

#### СПИСОК ЛИТЕРАТУРЫ

1. Volkov E.N., Averkin A.N. Explainable Artificial Intelligence in Medical Image Analysis: State of the Art and Prospects // XXVI Intern. Conf. on Soft Computing and Measurements (SCM). IEEE, 2023. P. 134–137. <https://doi.org/10.1109/SCM58628.2023.10159033>
2. Averkin A.N., Volkov E.N., Yarushev S.A. Possibilities of application of neuro-fuzzy networks for ophthalmologic image classification // Pattern Recognition Image Analysis. 2024. V. 34. № 3. P. 610–616. <https://doi.org/10.1134/S1054661824700421>
3. Averkin A.N., Volkov E.N., Yarushev S.A. Explainable artificial intelligence in deep learning neural nets-based digital images analysis // J. Comp. Systems Sci. Int. 2024. V. 63. № 1. P. 175–203. <https://doi.org/10.1134/S1064230724700138>
4. Рыжов А.П. О качестве классификации объектов на основе нечетких правил // Интеллектуальные системы. 2005. Т. 9. С. 253–264.
5. Krzywicki T., Brona P., Zbrzezny A.B. et al. A global review of publicly available datasets containing fundus images: characteristics, barriers to access, usability, and generalizability // J. Clin. Med. 2023. V. 12. № 10. P. 3587. <https://doi.org/10.3390/jcm12103587>
6. Jha D., Smedsrud P.H., Riegler M.A. et al. Resunet++: an advanced architecture for medical image segmentation // IEEE Intern. Sympos. Multimedia (ISM). 2019. P. 225–2255.
7. Van der Velden B.H.M., Kuijff B.H., Gilhuijs H.J. et al. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis // Med. Image Analysis. 2022. V. 79. P. 102470. <https://doi.org/10.1016/j.media.2022.102470>
8. Qian J., Li H., Wang J. et al. Recent advances in explainable artificial intelligence for magnetic resonance imaging // Diagnostics. 2023. V. 13. № 9. P. 1571. <https://doi.org/10.3390/diagnostics13091571>
9. Volkov E.N., Averkin A.N. Possibilities of explainable artificial intelligence for glaucoma detection using the LIME method as an example // XXVI Intern. Conf. on Soft Computing and Measurements (SCM). IEEE: Saint-Petersburg, 2023. P. 130–133. <https://doi.org/10.1109/SCM58628.2023.10159038>
10. Saeed W., Omlin C. Explainable Ai (Xai): a systematic meta-survey of current challenges and future opportunities // Knowledge-Based Systems. 2023. V. 263. P. 110273. <https://doi.org/10.1016/j.knosys.2023.110273>
11. Clement T., Kemmerzell N., Abdelaal M. et al. XAIR: a systematic metareview of explainable AI (XAI) aligned to the software development process // Mach. Lear. Knowledge Extraction. 2023. V. 5. № 1. P. 78–108. <https://doi.org/10.3390/make5010006>
12. Selvaraju R.R., Cogswell M., Das A. et al. Grad-cam: visual explanations from deep networks via gradient-based localization // Proc. IEEE Intern. Conf. on Computer Vision. Venice, 2017. P. 618–626.
13. Zhou B., Khosla A., Lapedriza A. et al. Learning deep features for discriminative localization // Proc. IEEE Conf. on Computer Vision and Pattern Recognition. Las Vegas, 2016. P. 2921–2929.
14. Cheng B., Girshick R., Dollar P. et al. Boundary IoU: improving object-centric image segmentation evaluation // Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Nashville, USA. 2021. P. 15334–15342.
15. Zhao R., Qian B., Zhang X. et al. Rethinking dice loss for medical image segmentation // IEEE Intern. Conf. on Data Mining (ICDM). Sorrento, Italy. IEEE, 2020. P. 851–860. <https://doi.org/10.1109/ICDM50108.2020.00094>
16. Hehn T., Kooij J., Gavrilu D. Fast and compact image segmentation using instance stixels // IEEE Transactions on Intelligent Vehicles. 2021. V. 7. № 1. P. 45–56. <https://doi.org/10.1109/TIV.2021.3067223>