

# Применение искусственного интеллекта в офтальмологии на примере решения задачи семантической сегментации изображения глазного дна

Н.С. Демин<sup>1,2</sup>, Н.Ю. Ильясова<sup>1,2</sup>, Р.А. Парингер<sup>1,2</sup>, Д.В. Кириш<sup>1,2</sup>

<sup>1</sup> ИСОИ РАН – филиал ФНИЦ «Кристаллография и фотоника» РАН,  
443001, Россия, г. Самара, ул. Молодогвардейская, д. 151;

<sup>2</sup> Самарский национальный исследовательский университет имени академика С.П. Королёва,  
443086, Россия, г. Самара, Московское шоссе, д. 34

## Аннотация

В данной работе представлены основные аспекты применения искусственного интеллекта в офтальмологии для диагностики и лечения глазных заболеваний на примере задачи семантической сегментации изображений глазного дна. Проведено сравнение классического подхода к семантической сегментации на основе текстурных признаков и предлагаемого подхода на основе нейронных сетей. Сформулированы основные проблемы применения нейросетевого подхода в биомедицине. Предложен способ выделения оптимальной зоны лазерного воздействия для проведения операции лазерной коагуляции на основе применения двух нейронных сетей. Первая сеть применялась для выделения анатомических объектов на глазном дне, а вторая – для выделения зоны макулярного отёка. Результат формировался из области отёка с учётом расположения на ней анатомических объектов. Был проведён сравнительный анализ нескольких архитектур нейронных сетей для решения задачи выделения области отёка. Лучшие результаты выделения зоны отёка показала нейронная сеть архитектуры Unet++.

**Ключевые слова:** изображение глазного дна, лазерная коагуляция, диабетическая ретинопатия, обработка изображений, сегментация, нейронная сеть, искусственный интеллект.

**Цитирование:** Демин, Н.С. Применение искусственного интеллекта в офтальмологии на примере решения задачи семантической сегментации изображения глазного дна / Н.С. Демин, Н.Ю. Ильясова, Р.А. Парингер, Д.В. Кириш // Компьютерная оптика. – 2023. – Т. 47, № 5. – С. 824-831. – DOI: 10.18287/2412-6179-CO-1283.

**Citation:** Demin NS, Ilyasova NY, Paringer RA, Kirsh DV. Application of artificial intelligence in ophthalmology for solving the problem of semantic segmentation of fundus images. Computer Optics 2023; 47(5): 824-831. DOI: 10.18287/2412-6179-CO-1283.

## Введение

В последнее время внедрение искусственного интеллекта, современных информационных технологий цифровой медицины в практику здравоохранения стремительно меняет способы диагностики и лечения, формы взаимодействия врачей с пациентами и коллегами, организацию лечения и восстановления здоровья [1, 2]. Все чаще роботизированные системы используются для поддержки диагностики и лечения заболеваний [3]. Согласно Прогнозу научно-технологического развития РФ на период до 2030 года, перспективными направлениями научных исследований являются проектирование интеллектуальных систем поддержки медицинских решений, а также оказание услуг по анализу медицинских данных [4].

Офтальмология остро нуждается в переходе на персонализированную медицину, которая позволила бы сделать качественный скачок в лечении глазных заболеваний [5]. Однако этот переход невозможен без разработки и внедрения принципиально новых интеллектуальных методов анализа биомедицинских данных пациентов.

Сахарный диабет признан одной из глобальных медико-социальных проблем современного общества.

Среди его наиболее тяжёлых и распространённых осложнений – диабетическая ретинопатия [6, 7, 8]. Это заболевание стало одной из основных причин нарушения зрения вплоть до необратимой слепоты. При своевременном лечении потерю зрения можно предотвратить более чем в 50 % случаев [9–13].

Одним из эффективных способов лечения диабетической ретинопатии является очаговая лазерная хирургия – нанесение множества дозированных микроскопических ожогов (коагулятов) в области возникновения макулярного отёка, вызванного поражением мелких кровеносных сосудов сетчатки глаза [14, 15, 16]. Эффективность такой процедуры зависит от опыта и квалификации конкретного хирурга-офтальмолога и, в первую очередь, от точности расстановки им коагулятов. При подготовке к операции специалист совмещает данные оптической когерентной томографии (ОКТ) и глазного дна пациента, на основании которых разрабатывает план лазерной коагуляции поражённых участков сетчатки. Однако ручная расстановка может быть весьма далёкой от оптимальной. Неравномерность расстановки точек прижигания создаёт либо риск увеличения травматичности в зонах избыточной коагуляции, либо снижает эффективность лечения на участках, где воздействие

было недостаточным. К тому же на планирование такой операции уходит много времени. Для повышения скорости подготовки к операции и улучшения равномерности расстановки зачастую используют стандартные шаблоны – паттерны из нескольких коагулятов, но добиться при этом соответствия индивидуальной форме отёка и расположению сосудов становится невозможно.

Современные системы в основном полагаются на использование предустановленного шаблона для создания карты фотокоагуляции [15, 16]. Из-за высокой изменчивости формы макулярного отёка и сосудистой системы единая карта фотокоагуляции не может быть реализована с использованием стандартной схемы [15, 16]. Однако офтальмологу предварительно необходимо проанализировать состояние сетчатки и глазного дна, чтобы убедиться, что фотокоагуляты наносятся в допустимых зонах. С одной стороны, этот метод обеспечивает более эффективную лазеркоагуляцию при правильно нанесённом паттерне, но с другой стороны, хирургу требуется дополнительное время для анализа состояния сетчатки [17].

Применение искусственного интеллекта позволит проводить высокоточное распознавание патологических и анатомических структур глазного дна с целью формирования зон лазерного воздействия и персонализированного плана оптимального расположения коагулятов в области диабетического макулярного отёка ДМО [18–21]. Фотокоагуляция будет проводиться исключительно в области поражённого участка глаза. Это увеличит качество лазерного лечения и объективной оценки объёма и локализации патологических структур, позволяющей прогнозировать результаты лечения и своевременно изменять тактику лечения диабетической ретинопатии ДР. По оценке врачей, внедрение такого подхода уменьшит вероятность лазерных ожогов за границами отёка в 9 раз, существенно сократит время подготовки пациента к операции и в целом снизит риск появления послеоперационных осложнений [16]. Оптимизированная лазеркоагуляция обеспечивает максимально равномерное воздействие лазерной энергии на пигментный эпителий поражённой сетчатки, что увеличивает безопасность лечения и его эффективность, а также позволяет избежать применения более дорогостоящих методов лечения [22, 23].

### **1. Различные подходы к семантической сегментации**

В основе интеллектуальной системы лазерной фотокоагуляции сетчатки лежат методы автоматизации принятия решений для лечения диабетической ретинопатии, позволяющие реализовать персонализированный подход к лечению диабетического макулярного отёка (ДМО). Для этого были использованы различные подходы семантической сегментации и новая методика нанесения коагулятов, учитывающая раз-

личные особенности расположения совокупности коагулятов с учётом индивидуальных особенностей анатомических структур и границ отёка у конкретного пациента. Для выделения анатомических и патологических зон исследовались и сравнивались два подхода семантической сегментации: текстурный и нейросетевой анализ.

Текстурный анализ является классическим способом сегментации, основанным на использовании текстурных признаков. Данные признаки хорошо изучены и успешно применяются к самым разнообразным объектам исследования [24–26]. Однако анализ признаков требует больших вычислительных ресурсов, что затрудняет их внедрение в медицинскую практику. Кроме того, текстурные признаки обладают недостаточной обобщаемостью для решения сложных задач.

В текстурном анализе сегментация выполняется в несколько этапов [24]:

1. Фрагментация изображения.
2. Вычисление текстурных признаков для каждого фрагмента [25] (этот этап является наиболее длительным: вычисление небольшого набора из пары десятков признаков для одного изображения размера  $1024 \times 1024$  пикселей может занимать несколько часов при использовании современного многоядерного процессора).
3. Классификация фрагментов на основе вычисленных значений текстурных признаков (на данном этапе производится классификация одного из пикселей фрагмента изображения по вектору значений вычисленных текстурных признаков, чаще всего классификатор строится с использованием деревьев решений).

Анализ полного набора текстурных признаков для сегментации изображений неэффективен, поэтому производится отбор признаков согласно индивидуальному критерию информативности дискриминантного анализа.

Второй подход использует методы машинного обучения на основе глубоких нейронных сетей.

Применение нейросетевого подхода для решения задачи семантической сегментации изображений глазного дна обусловлено рядом причин. Нейросетевые алгоритмы обладают хорошей обобщающей способностью; являются более точными (превзошли любые другие подходы) во множестве задач интеллектуального анализа данных; способны учитывать весь контекст изображения [27]. Однако их применимость ограничивается особенностями обучающего набора данных. Было выявлено, что применение нейронных сетей в биомедицине сопряжено с рядом специфических для этой области проблем:

1. Чаще всего из-за соблюдения политик конфиденциальности и трудоёмкости процедуры разметки, требующей высокой квалификации от специалиста, создать достаточно большой набор данных необходимого качества крайне затруд-

нительно [28]. Влияние проблемы недостаточного объёма данных можно нивелировать благодаря применению различных техник аугментации данных [29]. Так, в задачах, связанных с обработкой биомедицинских изображений, особенно действенной оказывается эластичная аугментация данных [30]: поворот на случайный угол, отражения, эластичная деформация.

2. Другая проблема, свойственная биомедицинским данным, – проблема выраженного дисбаланса классов, что является отличительной чертой [31]. Выборка содержит классы, которые на практике встречаются далеко не на всех изображениях сетчатки глаза, при этом относительная мощность этих классов чрезвычайно мала. Эту особенность необходимо учитывать при разработке алгоритма сегментации на основе нейронной сети. Решение задачи в случае, когда набор классов оказывается несбалансированным, является нетривиальным, однако существуют алгоритмы, которые нивелируют данную задачу [32, 33, 34].
3. Наиболее сложной проблемой является низкое качество маркировки данных [35], которое труднее всего выявить и устранить на этапе обучения алгоритма.

В ходе разработки интеллектуальной системы необходимо было предварительно решить целый ряд сопутствующих задач:

1. Исследовать применимость различных нейронных сетей для решения задачи семантической сегментации изображений глазного дна.
2. Определить факторы, которые необходимо учитывать, чтобы получить качественную сегментацию изображений глазного дна.
3. Определить оптимальную архитектуру и гиперпараметры нейронной сети.
4. Выявить особенности набора данных, которые необходимо учитывать при разработке алгоритмов нейронной сети.

Учитывая описанные выше особенности биомедицинских данных, для применения нейросетевых алгоритмов первоочередной задачей стала подготовка данных: нивелирование проблем несбалансированности и малого объёма выборки. Совместное использование различных техник аугментации данных позволило нам значительно расширить набор доступных изображений глазного дна, в результате чего мы получили более 5000 изображений. В ходе исследования для задачи семантической сегментации было адаптировано несколько различных архитектур нейронных сетей.

В работе [36] нами были представлены сравнительные исследования двух указанных подходов (текстурный и нейросетевой) к семантической сегментации. Исследования показали, что нейронные сети превосходят текстурные признаки по точности. Более

того, нейронные сети могут быть использованы для сегментации изображений глазного дна, которые были получены при различных условиях съёмки, в отличие от текстурных признаков. Использование предварительно обученных нейронных сетей и их дополнительное обучение на небольшом наборе данных вместе с использованием алгоритма балансировки и методик аугментации позволило разработать достаточно точные алгоритмы семантической сегментации изображений глазного дна на основе небольшого набора данных.

## **2. Нейросетевой подход для автоматического выделения оптимальной зоны лазерного воздействия**

Для формирования плана необходимо определять, какая зона может подвергаться обстрелу, а также какая зона является оптимальной, в том смысле, что за счёт исключения областей, нецелесообразных для воздействия, уменьшается воздействие на глазное дно, что должно способствовать повышению качества лечения.

Для формирования зоны лазерного воздействия использовались две нейронные сети. Первая сеть позволяет выделять анатомические объекты на изображениях глазного дна. Вторая сеть – выделяет зону отёка.

Алгоритм выделения оптимальной зоны лазерного воздействия состоит из 4 этапов:

1. С использованием первой нейронной сети на изображении выделяются анатомические объекты.
2. С использованием второй нейронной сети на изображении выделяется зона отёка.
3. Итоговые карты сегментации нейронной сети представляют собой вероятностное распределение, требующее постобработки. Поскольку данные карты выделяют область отёка или, говоря иначе, области лазерного воздействия, воздействие лазером на всю эту область – иррационально и опасно. В связи с этим по соображениям безопасности из данной области убираются места, соответствующие объектам глазного дна – диск зрительного нерва, фовеа, сосуды, экссудаты и лазеркоагуляты.
4. Результирующая же маска оптимальной зоны лазерного воздействия получается из предварительной маски путём её обработки с использованием морфологических функций эрозии и дилатации для уточнения границ макулярного отёка. А также обработки медианным фильтром для сглаживания неровностей маски.

### ***2.1. Архитектуры нейронных сетей***

Все используемые архитектуры нейронных сетей являются полносвёрточными нейронными сетями. Данный тип нейронных сетей лучше всего подходит для решения задач семантической сегментации, так

как учитывается весь контекст изображения, а сегментация выполняется за один проход.

В работе [36] сравнивались сети на основе Unet [37], где в качестве энкодера использовалась одна из сетей ResNet [38], DenseNet [39] или Xception [40]. Использовались предварительно обученные веса на наборе данных ImageNet [41], которые применялись для инициализации энкодера и фиксировались во время обучения. По результатам исследований было выявлено, что сеть XceptionUnet показывает наибольшую точность среди исследуемых. Эта сеть использовалась для первого пункта алгоритма, для выделения анатомических объектов на глазном дне.

Для поиска подходящей архитектуры нейронной сети для второго пункта алгоритма – выделения области лазерного воздействия – в настоящей работе применялись такие нейронные сети, как Unet, Unet++, MANet, Linknet, FPN, PSPNet, PAN, DeepLabV3.

Следует отметить, что все нейронные сети, используемые в рамках данной работы, основаны на сети Resnet-34, предобученной на изображениях из набора данных ImageNet. Использование предобученной нейронной сети позволяет быстрее обучать собственные сети, т.к. такие сети уже способны извлекать огромное число признаков из изображений.

Unet – одна из первых полносвёрточных архитектур нейронных сетей, успешно применённых для решения задачи семантической сегментации биомедицинских изображений. Архитектура Unet может быть представлена схематично в виде буквы U. Данная архитектура состоит из двух частей, из энкодера, выполняющего роль классификатора или экстрактора признаков, и декодера, который составляет карту сегментации изображения из признаков, выделенных энкодером. Данная сеть является классической в задачах сегментации биомедицинских изображений.

Unet++ – улучшенная версия архитектуры Unet, основным изменением в данной архитектуре является замена связующих между энкодером и декодером на небольшие промежуточные сети [42]. Данное изменение позволяет улучшить внутреннее представление признаков, что может повысить качество сегментации глазного дна.

Архитектура MANet нейронной сети была разработана для решения задачи семантической сегментации изображений, полученных в результате сканирования печени и головного мозга [43]. Особенностью данной нейронной сети является применение механизма внимания. Данный механизм моделирует человеческую способность фокусироваться на объекте или области интереса. Таким образом, данная особенность позволяет нейронной сети обращать внимание на определённые области, признаки, каналы в зависимости от рассматриваемой области, что может повысить точность выделения макулярного отёка.

Архитектура Linknet является аналогом архитектуры Unet, за исключением некоторых измене-

ний [44]. Так, в оригинальной Unet декодер и энкодер связаны между собой простыми связями. В архитектуре Linknet предлагается использовать residual connection, предложенные авторами архитектуры Resnet. Таким образом, выходы декодера складываются с выходами энкодера. Данное изменение позволяет бороться с затухающим градиентом, а также добавлять информацию к декодеру, что повышает качество сегментации.

Архитектура FPN также схожа с архитектурой Unet [45]. Её особенностью является то, как используются выходы каждого блока декодера. В Unet итоговой картой сегментации является выход из последнего слоя нейронной сети. В то время как в FPN итоговая карта сегментации составляется взвешенным суммированием выхода каждого блока декодера. Данная особенность позволяет лучше учитывать общий контекст изображения на разных масштабах, а также бороться с переобучением нейронной сети. Данная особенность полезна при малом количестве исходных данных.

Архитектура PSPNet кардинально отличается от рассмотренных ранее [46]. Здесь не применяется типичная U-образная архитектура. Так, PSPNet состоит из CNN, которая выполняет также роль энкодера признаков. Далее признаки с последнего слоя CNN попадают в специальный модуль, в котором к признакам применяются свёртки разных размеров. Данные свёртки выполняют локализацию признаков. Далее каждый выход свёртки увеличивается в размерах до размера исходного изображения с помощью слоя upsampling. После все слои конкатенируются вместе с картой признаков энкодера. По данному тензору применяют заключительный свёрточный слой, который формирует карту сегментации.

Архитектура PAN [47] аналогично Manet применяет механизм внимания для улучшения точности нейронной сети. В то время как в Manet применяют PAB (Position-wise Attention Block) и MFAB (Multi-scale Fusion Attention Block), в архитектуре PAN применяют FPA (Feature Pyramid Attention). Данный блок призван улучшить восприятие непосредственно блока внимания за счёт построения внутри FPA структуры, похожей на ту, что реализована в архитектуре сети FPN. Таким образом, блок FPA позволяет выделять признаки с разных масштабов, тем самым восполняя недостаток классических механизмов внимания.

Архитектура DeepLabV3 [48] также состоит из энкодера и декодера. Однако в отличие от всех предыдущих архитектур, основная идея данной архитектуры заключается в замене обычных свёрточных слоёв на слои разряженной свёртки (atrous convolution). Данные слои призваны помочь обучить нейронную сеть лучше понимать контекст изображения, что имеет особое значение, так как область отёка расположена в разных участках сетчатки на разных анатомических объектах.

### 3. Экспериментальные исследования

Для проведения экспериментов использовался набор из 50 изображений глазного дна, размеченных врачом-офтальмологом. На рис. 1 представлен пример изображения глазного дна и его разметка.

В силу малого размера исходного набора было необходимо существенно его расширить за счёт применения различных техник аугментации, с помощью которых было получено более 5000 изображений.

В данной работе применялись следующие техники аугментации изображений: отражение, поворот на случайный угол (от  $-30^\circ$  до  $30^\circ$ ), случайный сдвиг, эластичная трансформация. Применение аугментации позволяет успешно обучить нейронную сеть выпол-

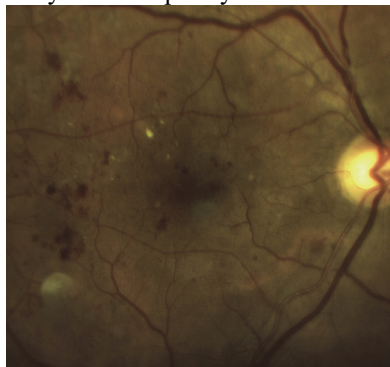


Рис. 1. а) Исходное изображение, б) бинарная маска разметки для макулярного отёка

нить семантическую сегментацию, а также бороться с переобучением сети при использовании небольших наборов данных.

Исходный набор данных был разбит на тренировочную и тестовую выборки в отношении 4 к 1. Таким образом, 40 изображений глазного дна составляли тренировочную выборку, а 10 изображений – тестовую. Все нейронные сети обучались при одинаковых параметрах:

- размер входа:  $512 \times 512 \times 3$ ;
- количество эпох: 150;
- функция ошибки: Cross Entropy [49];
- оптимизатор: Adam [50];
- скорость обучения: 0,001.

Табл. 1. Результаты обучения различных архитектур нейронных сетей

Архитектура	Значение метрики f1
Unet	<b>0,584</b>
Unet++	<b>0,562</b>
MAnet	0,508
Linknet	<b>0,575</b>
FPN	0,438
PSPNet	0,399
PAN	<b>0,512</b>
DeepLabV3	0,478

Лучшими по метрике стали сети: Unet, Unet++, Linknet, PAN. Дальнейшие исследования проводились с ними.

В ходе исследований было выявлено, что, если на вход нейронной сети подать, кроме изображения глазного дна, также карту объектов (оптический диск, макула, сосуды и т.д.), это позволит повысить точность модели при выделении области отёка. Объекты были выделены с помощью нейронной сети из первого пункта алгоритма выделения области лазерного воздействия зоны. Результаты соответствующего эксперимента представлены в табл. 2.

Как можно увидеть по результатам из таблицы, для трёх нейронных сетей повышается точность выделе-

ния области отёка. Дальнейшие исследования проводились с обучением нейронных сетей на изображениях глазного дна, совмещённых с картой объектов.

Табл. 2. Результаты эксперимента по проверке влияния добавления карты объектов

Архитектура	Без карты объектов	С картой объектов
Unet	0,584	<b>0,620</b>
Unet++	0,562	<b>0,632</b>
Linknet	0,575	<b>0,608</b>
PAN	<b>0,512</b>	0,446

Для изображений глазного дна на практике часто применяют особый вид предварительной обработки, при котором к исходному изображению глазного дна добавляется (взвешенная сумма) это же изображение, обработанное с помощью Гауссова фильтра с большим ядром. В табл. 3 представлены результаты эксперимента, проверяющего влияние данной предобработки на обучение нейронной сети в терминах метрики f1.

Табл. 3. Результаты эксперимента по проверке влияния предобработки изображений глазного дна

Архитектура	Без предобработки	С предобработкой
Unet	<b>0,620</b>	0,609
Unet++	<b>0,632</b>	0,626
Linknet	<b>0,608</b>	0,594
PAN	0,446	<b>0,468</b>

Результаты проведённых исследований показали, что для анализируемого набора данных применение

такой предварительной обработки не оказывает явного положительного эффекта на точность работы нейронной сети.

Для того, чтобы получить достоверные результаты эксперимента, была проведена k-fold кросс-валидация. В табл. 4 представлен усреднённый результат проведённого эксперимента в терминах метрик precision, recall, f1. Для оценки каждой группы бралась лучшая эпоха по метрике f1.

Табл. 4. Результаты кросс-валидации

Архитектура	precision	recall	f1
Unet	0,648	0,470	0,544
Unet++	0,716	<b>0,592</b>	<b>0,647</b>
Linknet	<b>0,728</b>	0,565	0,634
PAN	0,674	0,533	0,594

Таким образом, по полученным результатам, наиболее точной моделью оказалась Unet++. На рис. 2 представлен пример результатов семантической сегментации изображений.

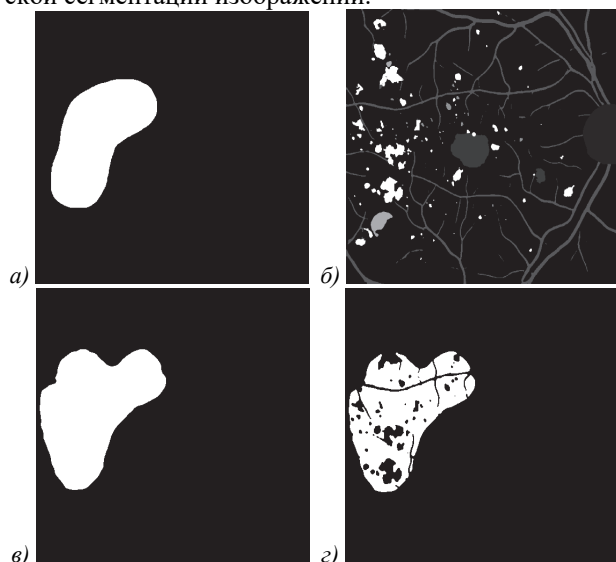


Рис. 2. Результаты предсказаний области макулярного отёка нейронной сетью: (а) экспертная разметка, (б) предсказание нейронной сети карты объектов, (в) результаты предсказаний нейронной сети области отёка, (г) обработанный результат

### Заключение

В статье представлены основные аспекты применения искусственного интеллекта в офтальмологии для диагностики и лечения глазных заболеваний на примере решения задачи семантической сегментации изображений глазного дна. Описаны подходы к сегментации биомедицинских изображений. Сформулированы четыре проблемы применения нейросетевого подхода к семантической сегментации. В результате проведённых исследований применения нейросетей было показано, что применение модели Unet++ для выделения области отёка даёт наиболее высокую точность относительно других рассмотренных в рамках работы архитектур. Итоговые карты сегментации

нейронной сети представляют собой вероятностное распределение, требующее постобработки. Зона отёка корректируется с учётом расположенных на ней анатомических ориентиров (маски анатомических ориентиров вычитаются из маски зоны отёка), таким образом получается предварительная маска зоны лазерного воздействия. Данные объекты (диск зрительного нерва, фовеа, сосуды, экссудаты и лазеркоагуляты) могут быть найдены с применением нейронной сети архитектуры XceptionUnet, выделяющей эти объекты на изображении глазного дна. Также отдельным этапом постобработки является применение морфологических функций для уточнения границ отёка. Конечным этапом постобработки является применение медианного фильтра для того, чтобы сгладить неровности итоговой маски.

### Благодарности

Работа выполнена при финансовой поддержке РФФИ (грант № 19-29-01135), в рамках выполнения работ по государственному заданию ФНИЦ «Кристаллография и фотоника» РАН.

### References

- [1] Rottier JB. Artificial intelligence: reinforcing the place of humans in our healthcare system. *La Revue du Praticien* 2018; 68(10): 1150-1151.
- [2] Fourcade A, Khonsari RH. Deep learning in medical image analysis: A third eye for doctors. *J Stomatol Oral Maxillofac Surg* 2019; 120(4): 279-288.
- [3] Gao A, et al. Progress in robotics for combating infectious diseases. *Sci Robot* 2021; 6(52): eabf1462.
- [4] Forecast of scientific and technological development of the Russian Federation until 2030. Source: <<http://static.government.ru/media/files/41d4b737638b91da2184.pdf>>.
- [5] Trinh M, Ghassibi M, Lieberman R. Artificial intelligence in retina. *Adv Ophthalmol Optom* 2021; 6: 175-185.
- [6] Vorobieva IV, Merkusheva DA. Diabetic retinopathy in patients with type 2 Diabetes Mellitus. *Epidemiology, a modern view of pathogenesis. Ophthalmology* 2012; 9(4): 18-21.
- [7] Dedov II, Shestakova MV, Galstyan GR. Prevalence of type 2 Diabetes Mellitus in the adult population of Russia (NATION study). *Diabetes Mellit* 2016; 19(2): 104-112.
- [8] Tan GS, Cheung N, Simo R. Diabetic macular edema. *Lancet Diab Endoc* 2017; 5: 143-155.
- [9] Amirov AN, Abdulaeva EA, Minkhuzina EL. Diabetic macular edema: Epidemiology, pathogenesis, diagnosis, clinical presentation, and treatment. *Kazan Medical Journal* 2015; 96(1): 70-74.
- [10] Doga AV, Kachalina GF, Pedanova EK, Buryakov DA. Modern diagnostic and treatment aspects of diabetic macular edema. *Ophthalmology Diabetes* 2014; 4: 51-59.
- [11] Bratko GV, Chernykh VV, Sazonova OV. On early diagnostics and the occurrence rate of diabetic macular edema and identification of diabetes risk groups. *Siberian Scientific Medical Journal* 2015; 35(1): 33-36.
- [12] Wong TY, Liew G, Tapp RJ. Relation between fasting glucose and retinopathy for diagnosis of diabetes: three population-based cross-sectional studies. *Lancet* 2008; 371(9614): 736-743.

- [13] Acharya UR, Ng EY, Tan JH, Sree SV, Ng KH. An integrated index for the identification of diabetic retinopathy stages using texture parameters. *J Med Syst* 2012; 36(3): 2011-2020.
- [14] Astakhov YuS, Shadrichev FE, Krasavina MI, Grigorieva NN. Modern approaches to the treatment of diabetic macular edema. *Ophthalmological Statements* 2009; 4: 59-69.
- [15] Zamytsky EA, Zolotarev AV, Karlova EV, Zamytsky PA. Analysis of the coagulates intensity in laser treatment of diabetic macular edema in a Navilas robotic laser system. *Saratov Journal of Medical Scientific Research* 2017; 13(2): 375-378.
- [16] Zamytskiy EA, Zolotarev AV, Karlova EV. Comparative quantitative assessment of the placement and intensity of laser spots for treating diabetic macular edema. *Russian Journal of Clinical Ophthalmology* 2021; 21(2): 58-62.
- [17] Kotsur TV, Izmailov AS. The effectiveness of laser coagulation in the macula and high-density microphotocoagulation in the treatment of diabetic maculopathy. *Ophthalmological Statements* 2016; 9(4): 43-45.
- [18] Whiting DR, Guariguata L, Weil C, Shaw J. IDF diabetes atlas: Global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Res Clin Pract* 2011; 94: 311-321.
- [19] Shirokanev AS, Kirsh DV, Ilyasova NYu, Kupriyanov AV. Investigation of algorithms for coagulate arrangement in fundus images. *Computer Optics* 2018; 42(4): 712-721. DOI: 10.18287/2412-6179-2018-42-4-712-721.
- [20] Ilyasova NYu, Demin NS, Shirokanev AS, Kupriyanov AV, Zamytskiy EA. Method for selection macular edema region using optical coherence tomography data. *Computer Optics* 2020; 44(2): 250-258. DOI: 10.18287/2412-6179-CO-691.
- [21] Ilyasova NYu, Shirokanev AS, Kupriyanov AV, Paringer RA. Technology of intellectual feature selection for a system of automatic formation of a coagulate plan on retina. *Computer Optics* 2019; 43(2): 304-315. DOI: 10.18287/2412-6179-2019-43-2-304-315.
- [22] Kozak I, Luttrull J. Modern retinal laser therapy. *Saudi J Ophthalmol* 2014; 29(2): 137-146.
- [23] Chhablani J, Mathai A, Rani P, Gupta V, Arevalo JF, Kozak I. Comparison of conventional pattern and novel navigated panretinal photocoagulation in proliferative diabetic retinopathy. *Investig Ophthalmol Vis Sci* 2014; 55: 3432-3438.
- [24] Ilyasova N, Paringer R, Kupriyanov A, Kirsh D. Intelligent feature selection technique for segmentation of fundus images. 2017 Seventh Int Conf on Innovative Computing Technology (INTECH) 2017: 138-143. DOI: 10.1109/INTECH.2017.8102433.
- [25] MaZda Web Site. Source: <<http://www.eletel.p.lodz.pl/programy/mazda/index.php>>.
- [26] Wu J, Poehlman S, Noseworthy MD, V. Kamath M. Texture feature based automated seeded region growing in abdominal MRI segmentation. *J Biomed Sci Eng* 2009; 02(01): 263-267. DOI:10.1109/BMEI.2008.352.
- [27] Gabbasov R, Paringer R. Influence of the receptive field size on accuracy and performance of a Convolutional Neural Network. 2020 Int Conf on Information Technology and Nanotechnology (ITNT) 2020: 1-4. DOI: 10.1109/ITNT49337.2020.9253219.
- [28] Arellano AM, Dai W, Wang S, Jiang X, Ohno-Machado L. Privacy policy and technology in biomedical data science. *Annu Rev Biomed Data Sci* 2018; 1: 115-129.
- [29] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for Deep Learning. *J Big Data* 2019; 6(1): 60.
- [30] Castro E, Cardoso JS, Pereira JC. Elastic deformations for data augmentation in breast cancer mass detection. 2018 IEEE EMBS Int Conf on Biomedical and Health Informatics (BHI) 2018: 230-344.
- [31] Ishwaran H, O'Brien R. Commentary: The problem of class imbalance in biomedical data. *J Thorac Cardiovasc Surg* 2021; 161(6): 1940-1941.
- [32] Charfe F, Rivera AJ, del Jesus MJ, Herrera F. MLSTMOTE: Approaching imbalanced multilabel learning through synthetic instance generation. *Knowl Based Syst* 2015; 89: 385-397.
- [33] Pereira RM, Costa YMG, Silla Jr. CN. MLTL: A multilabel approach for the Tomek link undersampling algorithm. *Neurocomputing* 2020; 383: 95-105.
- [34] Mukhin A, Kilbas I, Paringer R, Ilyasova N. Application of the gradient descent for data balancing in diagnostic image analysis problems. 2020 Int Conf on Information Technology and Nanotechnology (ITNT) 2020: 1-4. DOI: 10.1109/ITNT49337.2020.9253278.
- [35] Hao D, Zhang L, Sumkin J, Mohamed A, Wu S. Inaccurate labels in weakly-supervised deep learning: Automatic identification and correction and their impact on classification performance. *IEEE J Biomed Health Inform* 2020; 24(9): 2701-2710.
- [36] Paringer RA, Mukhin AV, Ilyasova NY, Demin NS. Neural network application for semantic segmentation of fundus. *Computer Optics* 2022; 46(4): 596-602. DOI: 10.18287/2412-6179-CO-1010.
- [37] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Int Conf on Medical Image Computing and Computer-Assisted Intervention* 2015: 234-241.
- [38] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proc IEEE Conf on Computer Vision and Pattern Recognition* 2016: 770-778.
- [39] Iandola F, Moskewicz M, Karayev S, Girshick R, Darrell T, Keutzer K. Densenet: Implementing efficient convnet descriptor pyramids. *arXiv Preprint*. 2014. Source: <<https://arxiv.org/abs/1404.1869>>.
- [40] Chollet F. Xception: Deep learning with depthwise separable convolutions. *Proc IEEE Conf on Computer Vision and Pattern Recognition* 2017: 1251-1258.
- [41] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 2012; 25: 1097-1105.
- [42] Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. UNet++: A nested U-Net architecture for medical image segmentation. In Book: Stoyanov D, Taylor Z, Carneiro G, Syeda-Mahmood T, Martel A, Maier-Hein L, Tavares JMRS, Bradley A, Papa JP, Belagiannis V, Nascimento JC, Lu Z, Conjeti S, Moradi M, Greenspan H, Madabhushi A, eds. *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Cham: Springer Nature Switzerland AG; 2018: 3-11.
- [43] Fan T, Wang G, Li Y, Wang H. MA-Net: A multi-scale attention network for liver and tumor segmentation. *IEEE Access* 2020; 8: 179656-179665.
- [44] Chaurasia A, Culurciello E. LinkNet: Exploiting encoder representations for efficient semantic segmentation. 2017 IEEE Visual Communications and Image Processing (VCIP) 2017: 1-4.
- [45] Lin T-Y, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2017: 936-944. DOI: 10.1109/cvpr.2017.106.
- [46] Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2017: 2881-2890.



- [47] Li H, Xiong P, An J, Wang L. Pyramid Attention Network for semantic segmentation. arXiv Preprint. 2018. Source: <<https://arxiv.org/abs/1805.10180>>.
- [48] Chen L-C, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for Semantic Image segmentation. arXiv Preprint. 2017. Source: <<https://arxiv.org/abs/1706.05587>>.
- [49] Ma Y-d, Liu Q, Quan Z-b. Automated image segmentation using improved PCNN model based on cross-entropy. Proc 2004 Int Symposium on Intelligent Multimedia, Video and Speech Processing 2004: 743-746.
- [50] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv Preprint. 2014. Source: <<https://arxiv.org/abs/1412.6980>>.

### Сведения об авторах

**Демин Никита Сергеевич**, 1994 года рождения, аспирант Самарского национального исследовательского университета имени академика С.П. Королева. Сфера научных интересов: интеллектуальный анализ медицинских изображений; цифровая обработка изображений; математическое моделирование; распознавание образов и искусственный интеллект. E-mail: [volfgunus@gmail.com](mailto:volfgunus@gmail.com).

**Ильясова Наталья Юрьевна**, 1966 года рождения. В 1991 году окончила с отличием Самарский государственный аэрокосмический университет имени С.П. Королева (СГАУ). В 1997 году защитила диссертацию на соискание степени кандидата технических наук, в 2015 году защитила диссертацию на соискание степени доктора технических наук. В настоящее время работает старшим научным сотрудником в Учреждении Российской академии наук Институте систем обработки изображений РАН – филиала ФНИЦ «Кристаллография и фотоника» РАН и одновременно доцентом кафедры технической кибернетики СГАУ. Круг научных интересов включает цифровую обработку сигналов и изображений, анализ и интерпретацию биомедицинских изображений. Имеет более 170 публикаций, три монографии (в соавторстве). E-mail: [ilyasova@ipsiras.ru](mailto:ilyasova@ipsiras.ru).

**Парингер Рустам Александрович**, 1990 года рождения, доцент кафедры технической кибернетики Самарского национального исследовательского университета имени академика С.П. Королева (Самарский университет), научный сотрудник Института систем обработки изображений РАН – филиала ФНИЦ «Кристаллография и фотоника» РАН. В 2013 году окончил факультет информатики СГАУ. Кандидат технических наук с 2017 года. Круг научных интересов включает интеллектуальный анализ данных, распознавание образов и искусственные нейронные сети. E-mail: [RusParinger@ssau.ru](mailto:RusParinger@ssau.ru).

**Кириш Дмитрий Викторович**, 1990 года рождения, окончил бакалавриат в 2012 году и магистратуру в 2014 году по направлению «Прикладная математика и информатика» в Самарском национальном исследовательском университете имени академика С.П. Королева (СГАУ). В 2019 году защитил диссертацию на соискание степени кандидата технических наук по специальности 05.13.17 «Теоретические основы информатики». В настоящий момент работает в должности доцента на кафедре технической кибернетики Самарского национального исследовательского университета имени академика С.П. Королева (Самарский университет), а также в должности научного сотрудника в Институте систем обработки изображений РАН – филиале ФНИЦ «Кристаллография и фотоника» РАН. Сфера научных интересов: цифровая обработка изображений и распознавание образов, машинное обучение, методы идентификации в кристаллографии. E-mail: [kirsh@ssau.ru](mailto:kirsh@ssau.ru).

ГРНТИ: 28.23.15

Поступила в редакцию 27 января 2023 г. Окончательный вариант – 29 мая 2023 г.



---

# Application of artificial intelligence in ophthalmology for solving the problem of semantic segmentation of fundus images

N.S. Demin<sup>1,2</sup>, N.Y. Ilyasova<sup>1,2</sup>, R.A. Paringer<sup>1,2</sup>, D.V. Kirsh<sup>1,2</sup>

<sup>1</sup> IPSI RAS – Branch of the FSRC “Crystallography and Photonics” RAS,  
443001, Samara, Russia, Molodogvardeyskaya 151;

<sup>2</sup> Samara National Research University, 443086, Samara, Russia, Moskovskoye Shosse 34

## Abstract

The paper presents main aspects of the application of artificial intelligence in ophthalmology for the diagnosis and treatment of eye diseases, considering the problem of semantic segmentation of fundus images as an example. The classic approach to semantic segmentation on the basis of textural features is compared to the proposed approach based on neural networks. Basic problems of using the neural network approach in biomedicine are formulated. We propose a new method for selecting an optimal zone of laser exposure for laser coagulation based on two neural networks. The first network is used for detecting anatomical objects in the fundus and the second one is used for selecting the area of macular edema. The region of interest is formed from the edema area while taking into account the location of anatomical objects in it. A comparative analysis of several architectures of neural networks for solving the problem of selecting the edema area is carried out. The best results in the selection of the edema area are shown by the neural network architecture of Unet++.

**Keywords:** fundus image, laser coagulation, diabetic retinopathy, image processing, segmentation, neural network, artificial intelligence.

**Citation:** Demin NS, Ilyasova NY, Paringer, RA Kirsh DV. Application of artificial intelligence in ophthalmology for solving the problem of semantic segmentation of fundus images. *Computer Optics* 2023; 47(5): 824-831. DOI: 10.18287/2412-6179-CO-1283.

**Acknowledgements:** This work was funded by the Russian Foundation for Basic Research under RFBR grant # 19-29-01135 and the Ministry of Science and Higher Education of the Russian Federation within a government project of FSRC “Crystallography and Photonics” RAS.

---

## Authors' information

**Nikita Sergeevich Demin** (born 1994), graduated (2019) with a master's degree in Applied Mathematics and Informatics. At present he is a postgraduate student of Samara University. The area of interests includes digital image processing, mathematical modeling, pattern recognition and artificial intelligence. E-mail: [volfgunus@gmail.com](mailto:volfgunus@gmail.com).

**Nataly Yurievna Ilyasova** (born 1966), graduated with honors from S.P. Korolyov Samara State Aerospace University (SSAU) (1991). She received her PhD (1997) and DSc (2015) in Technical Sciences. At present, she is a senior researcher at the IPSI RAS – Branch of the FSRC “Crystallography and Photonics”, and holding a part-time position of Associate Professor at SSAU's Technical Cybernetics sub-department. The area of interests includes digital signals and image processing, pattern recognition and artificial intelligence, biomedical imaging and analysis. Her list of publications contains more than 170 scientific papers and 3 monographs published with coauthors. E-mail: [ilyasova@ipsiras.ru](mailto:ilyasova@ipsiras.ru).

**Rustam Alexandrovich Paringer** (b. 1990), received Master's degree in Applied Mathematics and Informatics from Samara State Aerospace University (2013). He received his PhD in 2017. Associate professor of Technical Cybernetics department of Samara National Research University and researcher of IPSI RAS – Branch of the FSRC "Crystallography and Photonics". Research interests: data mining, machine learning and artificial neural networks. E-mail: [RusParinger@ssau.ru](mailto:RusParinger@ssau.ru).

**Dmitriy Victorovich Kirsh** (b. 1990), completed Bachelors' (2012) and Master's (2014) degrees in Applied Mathematics and Informatics from Samara State Aerospace University. Received his Ph.D. degree in Computer Science and Computer Engineering (2019). Currently he works as an associate professor at the Technical Cybernetics department of Samara University and also a researcher at the IPSI RAS – Branch of the FSRC “Crystallography and Photonics” RAS. E-mail: [kirsh@ssau.ru](mailto:kirsh@ssau.ru).

---

Received January 27, 2023. The final version – May 29, 2023.

---