

На правах рукописи



**ЛЕБЕДЕВ АНТОН АЛЕКСАНДРОВИЧ**

**ИССЛЕДОВАНИЕ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ  
ОБНАРУЖЕНИЯ ОБЪЕКТОВ НА ВИДЕОИЗОБРАЖЕНИЯХ  
В МЕДИЦИНСКИХ СИСТЕМАХ ПРИКЛАДНОГО  
ТЕЛЕВИДЕНИЯ**

2.2.13. Радиотехника, в том числе системы  
и устройства телевидения

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени  
кандидата технических наук

Ярославль – 2022

Работа выполнена в Центре искусственного интеллекта и цифровой экономики  
ФГБОУ ВО «Ярославский государственный университет им. П.Г. Демидова»

Научный руководитель: **Хрящев Владимир Вячеславович**, кандидат технических наук, доцент, доцент кафедры цифровых технологий и машинного обучения ЯрГУ им. П.Г. Демидова.

Официальные оппоненты: **Кузьмин Андрей Викторович**, доктор технических наук, доцент, профессор кафедры информационно-вычислительных систем ФГБОУ ВО «Пензенский государственный университет», г. Пенза.

**Мотыко Александр Александрович**, кандидат технических наук, доцент кафедры телевидения и видеотехники ФГАОУ ВО «Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)», г. Санкт-Петербург.

Ведущая организация: ФГБОУ ВО «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», г. Рязань.

Защита диссертации состоится 21.09.2022 в 16-00 на заседании диссертационного совета 24.2.281.01 при Владимирском государственном университете имени Александра Григорьевича и Николая Григорьевича Столетовых по адресу: 600000, г. Владимир, ул. Горького, д. 87, ВлГУ, корп. 3, ауд. 301.

С диссертацией можно ознакомиться в библиотеке Владимирского государственного университета имени Александра Григорьевича и Николая Григорьевича Столетовых и на сайте <http://diss.vlsu.ru>

Автореферат разослан 24.06.2022 г.

Отзывы на автореферат, заверенные печатью, просим направлять по адресу: 600000, г. Владимир, ул. Горького, д. 87, ВлГУ, РТ и РС Самойлову А.Г.

Ученый секретарь диссертационного совета,  
доктор технических наук, профессор



А.Г. Самойлов

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы.** Цифровая обработка видеоизображений – одна из перспективных и быстро развивающихся областей науки и техники, находящая применение при решении все более широкого круга задач, решаемых в ходе человеческой жизнедеятельности. Из них наиболее практически востребованными в данной области на текущем этапе научно-технического развития являются задачи автоматического обнаружения (детектирования) и классификации объектов на видеоизображениях. Они встречаются в различных сферах применения систем прикладного телевидения (СПТ): в охранных системах для обнаружения движущихся объектов, в промышленном телевидении для аналитики технологических процессов, в медицинских системах для поиска и визуализации различных патологий. В большинстве случаев для решения таких задач наряду с алгоритмами цифровой обработки изображений и технического зрения используются методы и алгоритмы машинного обучения.

Развитие подходов к обнаружению объектов на изображениях можно разделить на два этапа: использование классических алгоритмов машинного обучения (на основе эмпирически выбираемых признаков) и использование методов глубокого обучения (на основе применения сверточных нейронных сетей – СНС). В свою очередь нейросетевые алгоритмы обнаружения объектов разделяются на две основные категории: «одноэтапные» и «двухэтапные». Современные одноэтапные алгоритмы достигают сравнимой с двухэтапными точности обнаружения и имеют лучшую производительность.

Разработка и исследование подобных нейросетевых систем, анализирующих видеопоток в режиме реального времени, представляет собой актуальную научно-техническую проблему в области построения СПТ. Помимо задачи выбора оптимальной архитектуры нейронной сети и ее обучения исследователи сталкиваются с рядом проблем, касающихся подготовки, разметки и предобработки баз изображений, настройки алгоритмов аугментации (размножения) двумерных данных, а также алгоритмов постобработки, учитывающих специфику телевизионного сигнала.

Методы цифровой обработки и анализа видеоизображений все чаще применяются в медицинских СПТ, используемых для диагностики различных заболеваний. Накопление размеченных баз медицинских изображений и

видеоданных способствует появлению новых перспективных технологий для частичной автоматизации диагностики и контроля качества различных процедур. Одной из таких высокотехнологичных СПТ является эндоскопическая система, используемая, в частности, для колоноскопических исследований кишечника.

Несмотря на все более совершенное устройство современных эндоскопов, для получаемых с них видеоизображений характерен ряд особенностей, усложняющих их автоматический анализ:

- в силу конструктивных особенностей в видеоэндоскопах отсутствует автофокусировка, что приводит к отсутствию фокуса на объекте интереса;
- при движении видеоэндоскопа в полости органа часто возникает размытие изображений;
- в ПЗС- или КМОП-матрице присутствует тепловой шум;
- на слизистой часто можно наблюдать блики, вызванные отражением света, испускаемого источником света видеоэндоскопа;
- яркость и контрастность получаемого эндоскопического изображения может сильно меняться в зависимости от геометрии исследуемого органа;
- на выходе видеопроцессора эндоскопической системы обычно используется стандарт передачи видео с чересстрочной разверткой, что приводит к наличию на изображениях специфических искажений.

Поиск и удаление предраковых аномалий – один из наиболее эффективных способов борьбы с колоректальным раком. Особую значимость для анализа видеоизображений имеют полипы толстого кишечника, так как именно они могут перерасти в рак на последующей стадии. Риск появления и развития заболевания в значительной степени зависит от способности врача-эндоскописта обнаруживать полипы во время колоноскопической процедуры. Исследования показывают, что в среднем при обследовании кишечника врачом пропускается порядка 20% полипов. Современные эндоскопические устройства и диагностические инструменты повышают процент обнаружения полипов. Тем не менее, проблема обнаружения полипов остается важнейшей задачей для диагностики онкологии кишечника.

Таким образом, анализ российской и зарубежной научно-технической литературы показывает, что актуальной задачей на сегодняшнем этапе развития является применение алгоритмов цифровой обработки изображений и машинного обучения для частичной автоматизации диагностики в эндоскопических СПТ.

Это позволит повысить точность диагностики, уменьшить влияние человеческого фактора на качество исследований, снизить стоимость и временные затраты на их проведение.

Важнейший вклад в решение задач цифровой обработки изображений и распознавания образов внесли отечественные ученые В.П. Дворкович, А.В. Дворкович, А.И. Галушкин, Ю.И. Журавлев, В.А. Сойфер, В.Н. Безруков, Ю.В. Визильтер, А.С. Крылов, А.С. Конушин, Д.С. Ватолин, М.К. Чобану, Ю.С. Бехтин, М.Н. Фаворская, Ю.С. Радченко, Е.В. Медведева, А.Л. Приоров, С.С. Садыков и др. Важнейшие результаты в области обработки цифровых изображений с использованием глубоких нейронных сетей получены также следующими зарубежными учеными: Y. LeCun, Y. Bengio, G. Hinton, I. Goodfellow, K. He, R. Girshick, T.-Y. Lin, J. Redmon.

В области применения алгоритмов цифровой обработки для анализа эндоскопических видеоизображений следует отметить работы Н.А. Обуховой, А.А. Мотыко, С.В. Кашина, Р.О. Куваева, Д.В. Завьялова, R. Haidry, M. Hafner, D. Jha, Y. Shin, S. Ali, N.K. Tomar, H. Johansen, D. Johansen, J. Rittscher, и др.

**Целью работы** является повышение качества работы систем прикладного телевидения в эндоскопии путем разработки нейросетевых алгоритмов обнаружения объектов на видеоизображениях.

**Объектом исследования** являются нейросетевые алгоритмы обнаружения объектов на видеоизображениях в медицинских системах прикладного телевидения.

**Предметом исследования** является разработка, модификация и улучшение качества алгоритмов цифровой обработки видеоизображений и глубокого машинного обучения, применяемых для обнаружения патологических объектов в системах прикладного телевидения в эндоскопии.

**Для достижения указанной цели в диссертационной работе решаются следующие задачи:**

- разработка и исследование нейросетевых алгоритмов обнаружения объектов на видеоизображениях эндоскопических исследований;
- разработка методики полуавтоматического создания размеченной базы эндоскопических видеоданных для обучения нейросетевых моделей;
- усовершенствование алгоритмов предобработки и постобработки видеоизображений, учитывающих специфику видеозэндоскопических данных;

- тестирование разработанного алгоритма для анализа эндоскопического видеопотока в реальной медицинской системе прикладного телевидения.

**Методы исследования.** При решении поставленных задач использовались современные методы цифровой обработки сигналов и изображений, теория нейронных сетей и глубокого обучения. Для практической реализации алгоритмов применялись современные методы программирования и компьютерного моделирования на языках Python и C++.

### **Научная новизна полученных результатов**

В рамках работы получены следующие новые научные результаты:

1. Предложен нейросетевой алгоритм обнаружения полипов на статических изображениях в базах эндоскопических исследований.
2. Предложен нейросетевой алгоритм обнаружения полипов на видеоизображениях эндоскопических исследований, работающий в режиме реального времени.
3. Разработана методика полуавтоматического создания размеченной базы эндоскопических видеоданных для обучения нейросетевых моделей.
4. Усовершенствованы с учетом специфики видеоэндоскопических данных алгоритмы аугментации и постобработки видеоизображений.

### **Практическая значимость полученных результатов**

1. Разработанный нейросетевой алгоритм позволяет обнаруживать полипы на видеопотоке с эндоскопической системы в режиме реального времени. Он может быть использован в медицинских системах прикладного телевидения для частичной автоматизации диагностики, а также для контроля, обучения и подготовки профильных врачей.
2. Значение метрики средней точности для разработанного алгоритма обнаружения полипов на эндоскопических изображениях с модифицированным алгоритмом аугментации данных составляет 92,5% на общедоступной базе статических изображений Kvasir-SEG, что позволяет рекомендовать данный алгоритм для построения на его основе модуля анализа видеопотока в эндоскопической системе прикладного телевидения.
3. Предложена и апробирована на наборе из более чем 31 000 видеок кадров методика подготовки базы видеоданных эндоскопических исследований с разметкой областей интереса для обучения нейросетевых моделей, позволяющая сократить соответствующие трудозатраты на 30-40%.

4. Установлено, что применение предложенного алгоритма постобработки позволяет улучшить качество системы обнаружения полипов на видеопотоке в среднем на 1,5-2%.

Результаты работы внедрены в клиническую практику эндоскопического отделения Ярославской областной клинической онкологической больницы, а также в перспективные разработки компаний ООО «Точка зрения» (г. Ярославль) и ООО «Смарт анализ» (г. Ярославль). Проведено полугодовое тестирование предложенных алгоритмов в рамках специализированного программно-аппаратного комплекса в ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России.

Отдельные результаты диссертационной работы внедрены в учебный процесс ЯрГУ им. П.Г. Демидова в рамках дисциплин «Системы технического зрения» и «Цифровая обработка изображений», а также в научно-исследовательские работы при выполнении гранта РФФИ № 19-37-90153 и договора с Фондом содействия инновациям № 3183ГС1/48626 от 26.08.2019. Получены два свидетельства о государственной регистрации программ для ЭВМ (№ 2020618382, № 2022610262) и свидетельство о государственной регистрации базы данных (№ 2022620355).

**Достоверность полученных научных результатов** диссертационной работы обеспечивается корректным использованием математического аппарата и полученными экспериментальными данными, подтверждающими теоретические гипотезы и согласующимися с результатами, известными из российских и зарубежных научно-технических источников.

**Апробация работы.** Результаты работы обсуждались на следующих научно-технических семинарах и конференциях:

1. 20-я – 23-я международные конференции «Цифровая обработка сигналов и ее применение», Москва, 2018–2021.
2. 17-я международная научно-техническая конференция «Новые информационные технологии и системы», Пенза, 2020.
3. 14-я международная научно-технической конференция «Перспективные технологии в средствах передачи информации», Владимир, 2021.
4. 12-я международная конференция по машинному зрению ICMV 2019 (International Conference on Machine Vision), Нидерланды, Амстердам, 2019.
5. 18-й и 19-й международные симпозиумы «IEEE East-West Design & Test Symposium», 2020-2021, Болгария, Варна / Грузия, Батуми.

6. 77-я научно-техническая конференция Санкт-Петербургского НТО РЭС им. А.С. Попова, посвященная Дню радио, Санкт-Петербург, 2022.
7. 77-я всероссийская конференция (с международным участием) «Радиоэлектронные устройства и системы для инфокоммуникационных технологий», Москва, 2022.

**Публикации.** По теме диссертации опубликовано 17 научных работ, из них 4 статьи в журналах из Перечня ВАК, 4 работы, индексируемые в базе SCOPUS, и 9 докладов на научно-технических конференциях.

**Структура и объем работы.** Диссертация состоит из введения, трех глав, заключения, списка литературы, содержащего 159 наименований, и трех приложений. Она изложена на 131 странице машинописного текста, содержит 35 рисунков и 9 таблиц.

#### **Основные научные положения и результаты, выносимые на защиту**

- алгоритм обнаружения объектов на статических изображениях эндоскопических исследований на основе сверточной нейронной сети с модифицированным алгоритмом аугментации данных, для которого значение метрики  $AP@[0,25..0,75]$  составило 92,5%, что превосходит известные аналоги на 5% и более;
- алгоритм обнаружения объектов на видеоизображениях эндоскопических исследований с усовершенствованной процедурой постобработки, который позволил улучшить значение метрики  $AP@0,50$  на базе видеоданных ЯОКОБ в среднем на 1,7%;
- методика полуавтоматического создания размеченной базы эндоскопических видеоданных для обучения нейросетевых моделей, позволяющая сократить соответствующие трудозатраты на 30-40%.

### **ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ**

**Во введении** обоснована актуальность выбранной темы, сформулированы цель и задачи исследования, изложены основные положения, выносимые на защиту, показаны научная новизна и практическая значимость работы.

**В первой главе** рассматриваются нейросетевые алгоритмы обнаружения объектов на изображениях, а также метрики для оценки качества подобного класса алгоритмов, такие как точность (Precision,  $P$ ), полнота или чувствительность (Recall,  $R$ ) и средняя точность (Average Precision,  $AP$ ):



$$P = \frac{TP}{TP + FP}, R = \frac{TP}{TP + FN}, AP = \frac{\sum_{r=1}^N P_r}{N},$$

где  $TP$  – число истинно положительных результатов,  $FP$  – число ложноположительных результатов,  $FN$  – число ложноотрицательных результатов,  $P_r$  – значение точности, соответствующее определенному значению полноты  $r$ , а  $N$  – множество различных значений полноты.

Для оценивания качества работы алгоритмов используются метрики средней точности, рассчитанные при различных значениях порога отношения пересечения к объединению (Intersection over Union,  $IoU$ ): 0,25, 0,50 и 0,75. Данные метрики будем обозначать как  $AP@0,25$ ,  $AP@0,50$  и  $AP@0,75$  соответственно. Для комплексной оценки по множеству порогов  $IoU$  используется метрика  $AP@[0,25..0,75]$ , которая рассчитывается как среднее 11-ти значений  $AP$ , рассчитанных при 11 различных порогах  $IoU$  в пределах от 0,25 до 0,75 с шагом 0,05.

Дополнительно используется метрика среднего гармонического точности и полноты ( $F1$ ):

$$F1 = 2 \cdot \frac{P \cdot R}{P + R} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}.$$

Рассматриваются прикладные телевизионные системы и их применение в области эндоскопических исследований. Из проведённого анализа литературы можно сделать вывод о том, что разработка систем анализа видеопотока в эндоскопической системе с применением современных СНС для обнаружения полипов представляет собой актуальную научно-техническую задачу.

**Во второй главе** исследуются нейросетевые алгоритмы обнаружения полипов на эндоскопических изображениях кишечника.

Для обучения и тестирования алгоритмов используется открытый набор изображений Kvasir-SEG, так как на данный момент из публично доступных баз он предлагает наибольшее количество размеченных примеров. Jha D. и др. в своей работе<sup>1</sup>, сравнивая различные архитектуры СНС для обнаружения полипов, делают вывод, что наиболее подходящим по соотношению точность/производительность является нейросетевой алгоритм на базе архитектуры YOLOv4. Однако авторы не

---

<sup>1</sup> Jha D., Ali S., Tomar N. K., Johansen H.D., Johansen D., Rittscher J., Riegler M.A., Halvorsen P. Real-Time Polyp Detection, Localization and Segmentation in Colonoscopy Using Deep Learning // Computer Vision and Pattern Recognition, 2021, pp. 40496 – 40510.

рассматривают популярный алгоритм, часто дающий хорошие результаты на различных задачах обнаружения объектов – SSD, а также недавно предложенную модификацию одноэтапного алгоритма – YOLOv5.

Архитектура сверточной нейронной сети SSD приведена на Рисунке 1. В качестве базовых сетей в работе выбраны сверточные нейронные сети VGG с размерами входного слоя  $300 \times 300$  и  $512 \times 512$ , а также MobileNet.

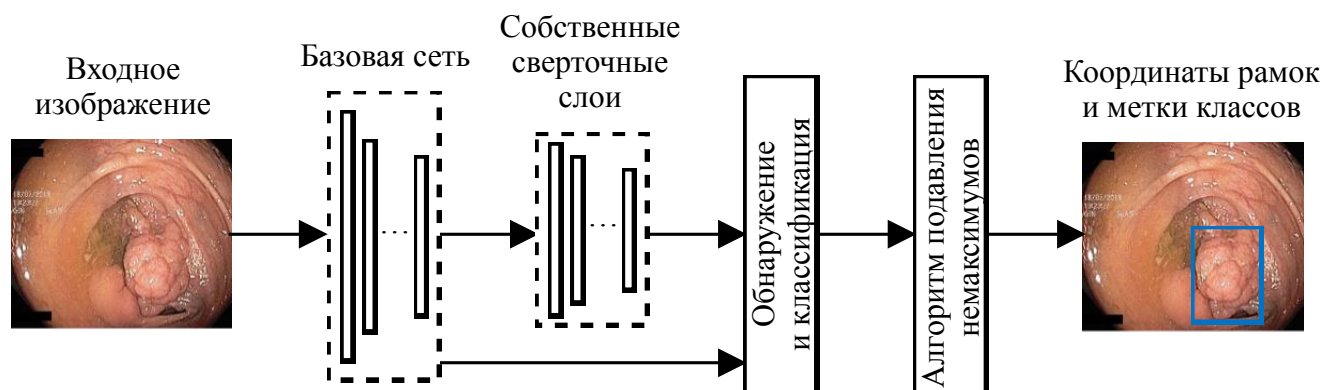


Рисунок 1 – Архитектура сверточной нейронной сети SSD

Архитектура сверточной нейронной сети YOLOv5 приведена на Рисунке 2. Существует 5 модификаций архитектуры YOLOv5 – N, S, M, L, X, отличающиеся количеством слоев и параметров сети. Модели образованы путем масштабирования базовой архитектуры по ширине и глубине.

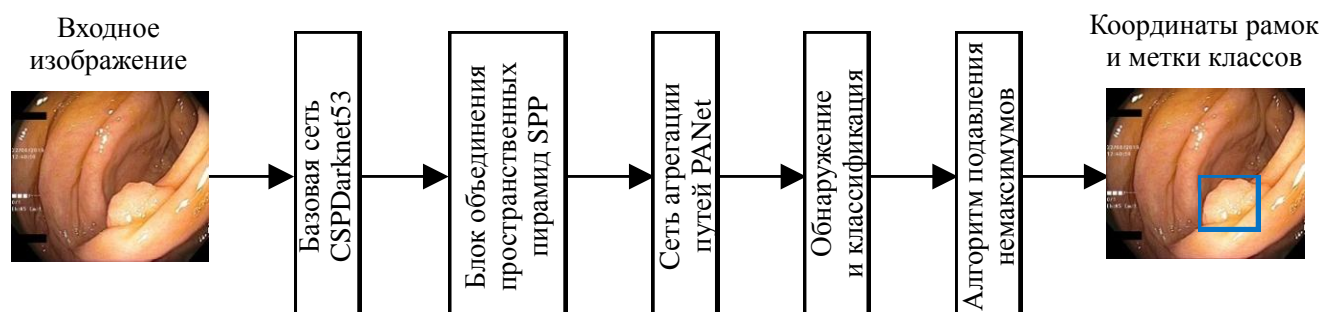


Рисунок 2 – Архитектура сверточной нейронной сети YOLOv5

Современные алгоритмы на основе СНС требуют большого количества обучающих примеров. Их получение посредством сбора и разметки изображений – рутинный процесс, требующий значительных ресурсов и трудозатрат как со стороны врачей, так и специалистов в области цифровой обработки изображений. Вместо этого во многих практических случаях целесообразно использовать аугментацию данных с помощью различных математических преобразований. Учитывая специфику эндоскопических изображений, в общепринятых стандартных методах аугментации можно использовать дополнительные преобразования, такие как: поворот изображений на случайное значение в пределах 180 градусов и

случайное с вероятностью 0,5 отзеркаливание изображения относительно горизонтальной оси. Используя предложенный модифицированный вариант аугментации данных, для алгоритма обнаружения полипов на основе сети SSD удалось улучшить метрику  $AP@0,50$  на 1,4%, а для алгоритма на основе сети YOLOv5 – на 3,0%.

Для решения поставленной задачи предложены и исследованы нейросетевые алгоритмы обнаружения объектов (HAOO) с модификацией процедуры аугментации. Три из них основаны на архитектуре SSD (HAOO<sub>SV300</sub>, HAOO<sub>SV512</sub>, HAOO<sub>SM300</sub>) и четыре – на архитектуре YOLOv5 (HAOO<sub>YS</sub>, HAOO<sub>YM</sub>, HAOO<sub>YL</sub>, HAOO<sub>YX</sub>).

Для сравнения с известными аналогами база эндоскопических изображений Kvasir-SEG разделялась на тестовый и обучающий набор: 120 изображений отбиралось для тестового набора и 880 – для обучающего. Случайным образом сформировано 10 обучающих и тестовых наборов, после чего каждый алгоритм обучался и тестировался на каждой из 10 пар, а в финальные результаты выбиралось усреднённое по 10 наборам значение метрики оценки качества работы алгоритмов.

Наилучший результат среди алгоритмов, основанных на архитектуре SSD, получен для алгоритма HAOO<sub>SV512</sub>, а для алгоритмов на основе YOLOv5 – для HAOO<sub>YL</sub>. На Рисунке 3 представлены кривые зависимости точности от полноты с указанием метрики  $AP@0,50$  для вышеперечисленных алгоритмов, обученных на первом из десяти случайных наборов.

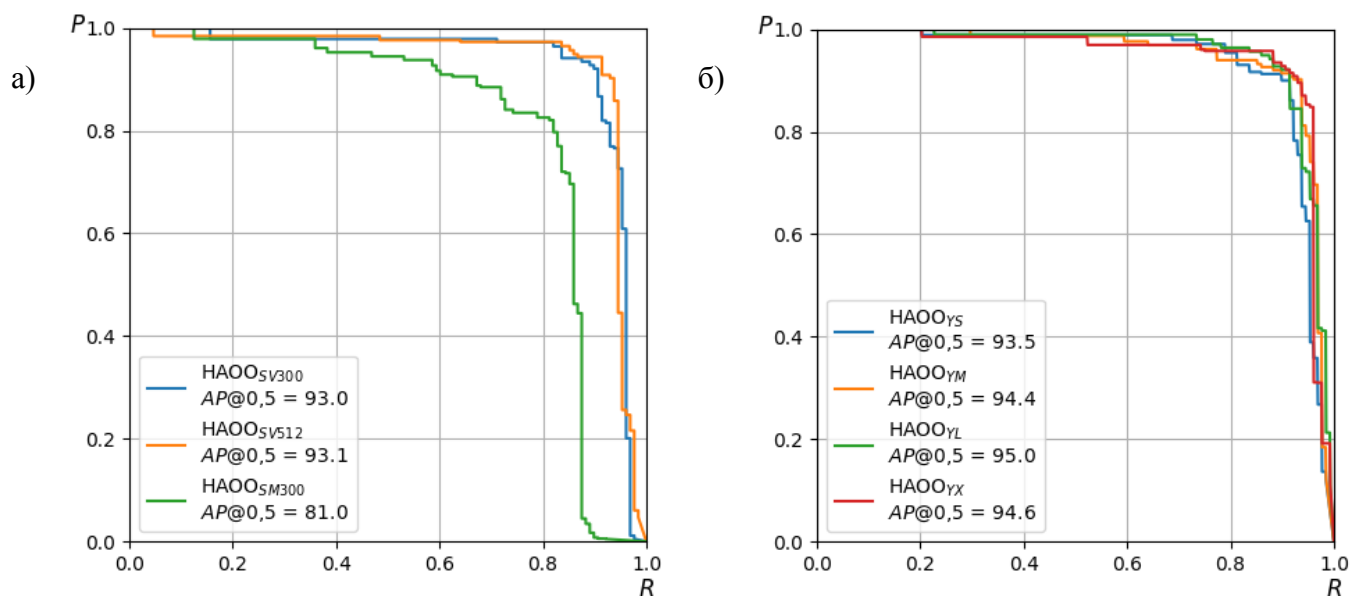


Рисунок 3 – Зависимости точности от полноты, полученные на базе изображений Kvasir-SEG:

а) для алгоритмов на основе SSD; б) для алгоритмов на основе YOLOv5

В Таблице 1 приведены метрики оценки качества для предложенных алгоритмов, а также результаты для известных аналогов.

Таблица 1

Сравнение результатов обнаружения полипов на базе Kvasir-SEG

Метрика	$AP@[0,25..0,75]$	$AP@0,25$	$AP@0,50$	$AP@0,75$	FPS
<b>Существующие алгоритмы</b>					
EfficientDet-D0	47,6	68,5	50,5	22,8	35,00
Faster R-CNN	78,7	89,5	84,2	56,6	8,00
RetinaNet50	87,0	93,9	90,9	69,7	16,20
RetinaNet101	<b>87,4</b>	<b>94,8</b>	<b>90,9</b>	71,3	16,80
YOLOv3+spp	81,0	88,6	85,3	<b>75,9</b>	45,01
YOLOv4	85,1	91,2	82,3	<b>75,9</b>	48,00
ColonSegNet	80,0	90,0	81,7	67,1	<b>180,00</b>
<b>Предложенные алгоритмы</b>					
HAOOSV300	89,3	96,3	93,3	70,0	71,97
HAOOSV512	90,2	96,6	93,8	74,9	46,33
HAOOSM300	76,4	90,6	82,9	44,6	<b>80,16</b>
HAOOSYS	88,6	94,0	90,8	73,8	79,36
HAOOSYM	91,8	95,9	93,3	82,8	61,35
<b>HAOOSYL</b>	<b>92,5</b>	<b>96,7</b>	<b>93,9</b>	<b>85,1</b>	54,05
HAOOSYX	91,5	95,6	92,3	84,6	43,85

Среди известных аналогов лучшие значения метрик средней точности при малых и средних порогах  $IoU$  достигает алгоритм RetinaNet101, а лучшее значение метрики оценки скорости работы алгоритма соответствует алгоритму ColonSegNet. Среди предложенных нейросетевых алгоритмов обнаружения лучшие результаты демонстрирует HAOOSYL. Стоит отметить, что алгоритм HAOOSV512 показал сопоставимый результат при низких и средних порогах  $IoU$ , однако при более высоких значениях  $IoU$  предпочтительней работают алгоритмы на базе архитектуры YOLOv5.

**В третьей главе** производится разработка и анализ алгоритма обнаружения полипов на эндоскопических видеоизображениях.

Во многих медицинских СПТ ставятся задачи обнаружения объектов на видеоданных. Это более сложная научно-техническая задача в сравнении с автоматическим анализом статических изображений, однако именно она и возникает в большинстве случаев в эндоскопической диагностике. В эндоскопических системах для передачи видеопотока чаще всего используются стандарт HD-SDI SMPTE 292M с разрешением 1080i и частотой обновления 50 Гц, что формирует высокие требования к задержкам, вносимым алгоритмами автоматического анализа видеопотока. Однако, используя метод квантования весов СНС, можно улучшить

производительность алгоритмов на их основе без значительного ухудшения качества их работы.

Применяя процедуру квантования параметров СНС до половинной точности представления чисел с плавающей точкой, удалось уменьшить среднее время обработки одного изображения для алгоритма НАОО<sub>УХ</sub> с 22,8 мс до 13,1 мс, что эквивалентно скорости обработки 76 кадров в секунду. При этом его средняя точность уменьшилась не более чем на 0,1%. Таким образом, модифицированный алгоритм НАОО<sub>УХ</sub> подходит для обработки видеопотока с эндоскопической системы в режиме реального времени.

Независимое применение разработанного алгоритма к каждому кадру видеопотока не использует корреляцию соседних кадров. Учесть эту особенность можно с помощью добавления в алгоритм обнаружения объектов этапа постобработки. При таком подходе выполняется связывание обнаружений между кадрами с помощью классификатора на основе логистической регрессии, чтобы различать, принадлежат ли два обнаружения одному и тому же экземпляру объекта или нет. Структурная схема всего алгоритма постобработки обнаружений на видеопотоке изображена на Рисунке 4. На первом этапе алгоритма происходит составление пар обнаружений, полученных с соседних кадров видеопотока, как показано на Рисунке 5.

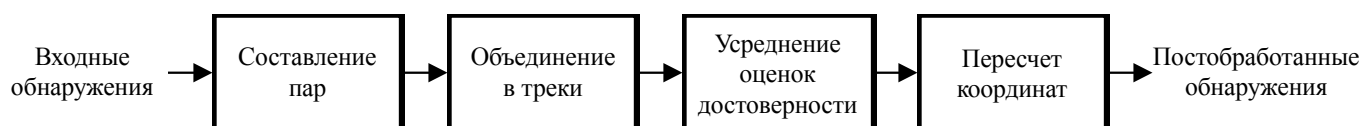


Рисунок 4 – Структурная схема алгоритма постобработки обнаружений на видеоданных

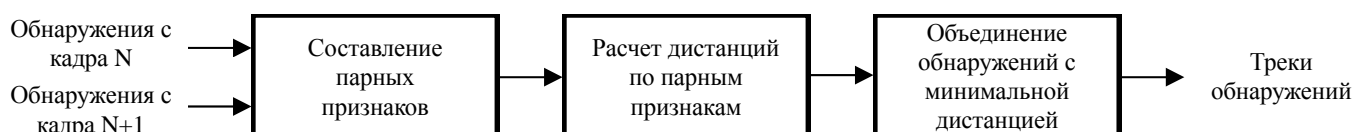


Рисунок 5 – Структурная схема алгоритма объединения обнаружений в треки

В качестве парных признаков, вычисляемых для каждой возможной пары обнаружений, используются следующие параметры:

$$f_{loc} = \{IoU, d_{centers}\}, f_{geo} = \{ratio_w, ratio_h\}, f_{sem} = cc_t^i \cdot cc_{t+1}^j,$$

где  $IoU$  – отношение площади пересечения между обнаружениями к площади объединения обнаружений,  $d_{centers}$  – евклидово расстояние между центрами обнаружений,  $ratio_w$  и  $ratio_h$  – соотношения ширин и высот обнаружений

соответственно,  $cc_t^i$  и  $cc_{t+1}^j$  – вектора оценок достоверности принадлежности к определенному классу объектов для рассматриваемой пары обнаружений в текущем и следующем кадрах, а  $f_{sem}$  – их скалярное произведение. Используя парные признаки, функцию подобия можно определить как:

$$L(d_t^i, d_{t+1}^j) = f_{sem} \cdot \varphi(f_{loc}, f_{geo}),$$

где  $\varphi$  – классификатор на основе логистической регрессии, предсказывающий, принадлежат ли два обнаружения одному и тому же экземпляру объекта, учитывая их парные признаки.

Определим также величину, обратную предложенной функции подобия, – дистанцию между двумя обнаружениями:

$$D(d_t^i, d_{t+1}^j) = 1 - L(d_t^i, d_{t+1}^j).$$

На основе матрицы дистанций пар обнаружений происходит связывание обнаружений в треки, что позволяет исправлять ошибочно классифицированные обнаружения или устранять неоднозначность с низким значением оценки достоверности принадлежности обнаружения к определенному классу.

Для обучения и тестирования алгоритмов, совместно с врачами эндоскопического отделения Ярославской областной клинической онкологической больницы собрана собственная база видеоданных (база ЯОКОБ). Обучающий набор состоит из 50 видеозаписей, содержащих 165 665 кадров, из которых 114 754 содержат полипы. Тестовый набор состоял из 9 видеозаписей, содержащих 11 192 кадра, из которых 7057 содержат полипы. Видеоданные в базе ЯОКОБ брались из архива, в котором видеоданные перекодированы с использованием видеокодека H.264, а их разрешение варьировалось от 626×532 до 760×436 пикселей.

На Рисунке 6 изображены зависимости точности от полноты для алгоритма  $HAOO_{YX}$  без применения этапа постобработки и с постобработкой (алгоритм  $HAOO_{YXVID}$ ), полученные на базе ЯОКОБ.

Алгоритм постобработки позволил улучшить качество обнаружения полипов на видеопоследовательностях, при этом значение метрики  $AP@0,50$  выросло с 81,6% до 83,3%. Кроме того, информация о треках полипов позволяет в дальнейшем реализовать автоматизированный подсчет числа полипов, осуществлять выбор лучшего кадра из трека для составления медицинского отчета, создавать аннотации для навигации по видеоданным исследования и т.д.

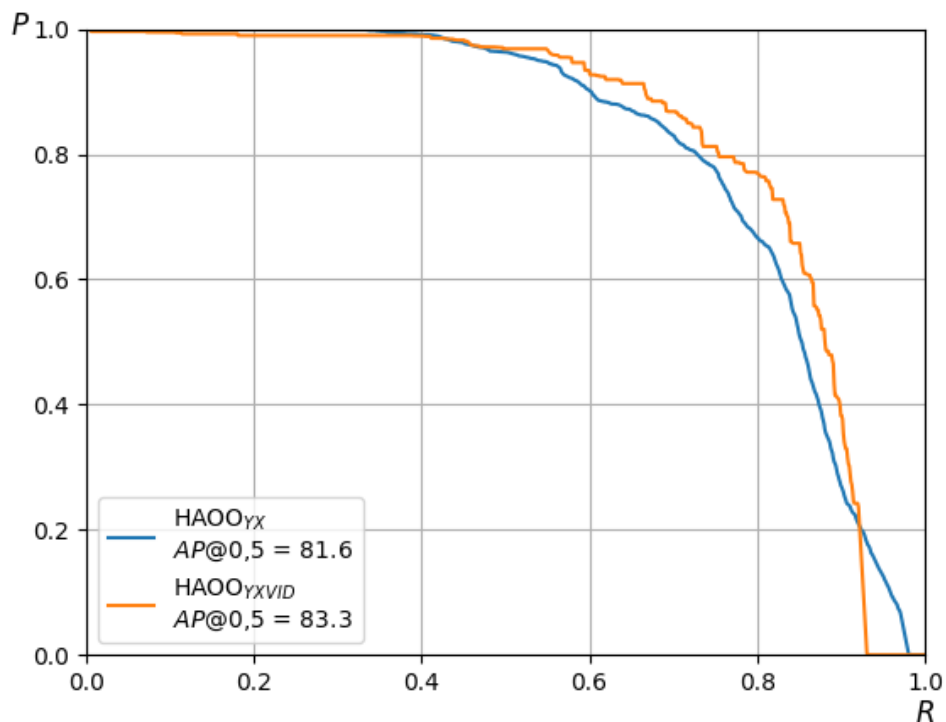


Рисунок 6 – Зависимости точности от полноты для алгоритмов HAOO<sub>YX</sub> и HAOO<sub>YXVID</sub>

Пример работы алгоритма постобработки на серии последовательных кадров приведен на Рисунке 7б. Для сравнения на Рисунке 7а приведены исходные результаты работы базового алгоритма обнаружения полипов. Видно, что недостающие обнаружения восстановлены за счет обнаружений из соседних кадров с помощью предложенного алгоритма постобработки.

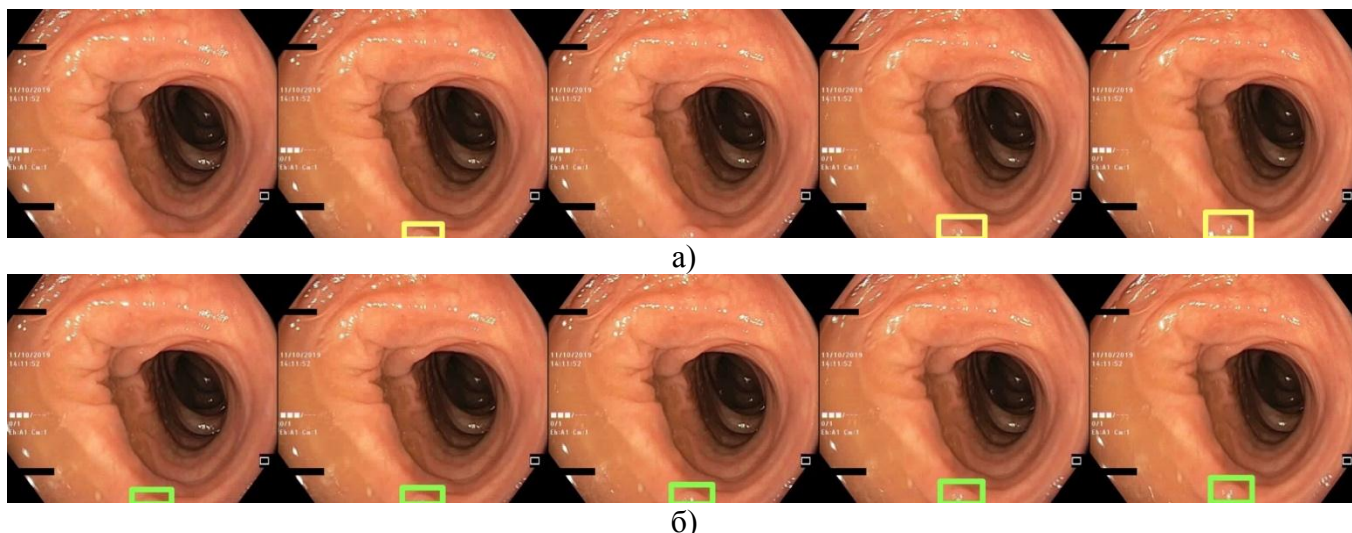


Рисунок 7 – Визуализация работы алгоритмов: а) HAOO<sub>YX</sub>; б) HAOO<sub>YXVID</sub>

Для сбора тестовой базы видеоданных в максимально практических условиях работы системы эндоскопического прикладного телевидения разработан специальный стенд (Рисунок 8), оборудованный платой видеозахвата. Он дополняет

эндоскопическую систему модулем анализа видеопотока и дополнительным монитором, на который выводится обработанный видеопоток.

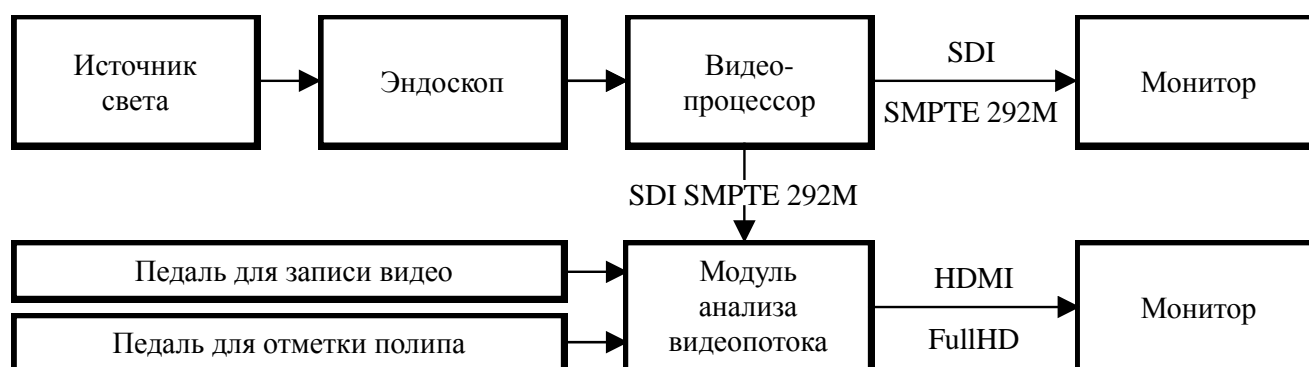


Рисунок 8 – Схема разработанного стенда с модулем анализа видеопотока

С помощью данного стенда в ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России (Пироговский центр, г. Москва) во временном интервале с сентября 2021 года по февраль 2022 года записаны 450 видеопоследовательностей колоноскопических исследований в условиях реальной клинической практики. Из них 207 записей содержали полипы, временные интервалы которых отмечались врачами-эндоскопистами с помощью педалей на тестовом стенде. Разрешение видеозаписей составляло  $1400 \times 1080$  пикселей (в оригинале FULL HD  $1920 \times 1080$  пикселей, но обрезалась неинформативная область кадров с персональными данными пациентов). Из них отобраны 43 наиболее информативных, которые разбивались на кадры, в каждом из которых отмечались местоположения полипов. Таким образом, собрана тестовая база видеофрагментов (ПРГВ), состоящая из 31 020 кадров и 27 133 экспертно размеченных рамок, локализирующих полипы в кадрах.

Для того чтобы снизить огромные трудозатраты эндоскопистов экспертного класса в процессе разметки видеофрагментов предложена методика полуавтоматической разметки областей интереса для последующего обучения и/или тестирования нейросетевых моделей. Данная методика предполагает формирование основы для разметки алгоритмом, обученным на предыдущих этапах исследования («алгоритм»), с доработкой полученной разметки специалистом по работе с изображениями («разметчик»), который, в свою очередь, консультируется с врачами-эндоскопистами («врач»). Структурная схема предложенной методики полуавтоматической разметки областей интереса на видеоданных приведена на Рисунке 9.



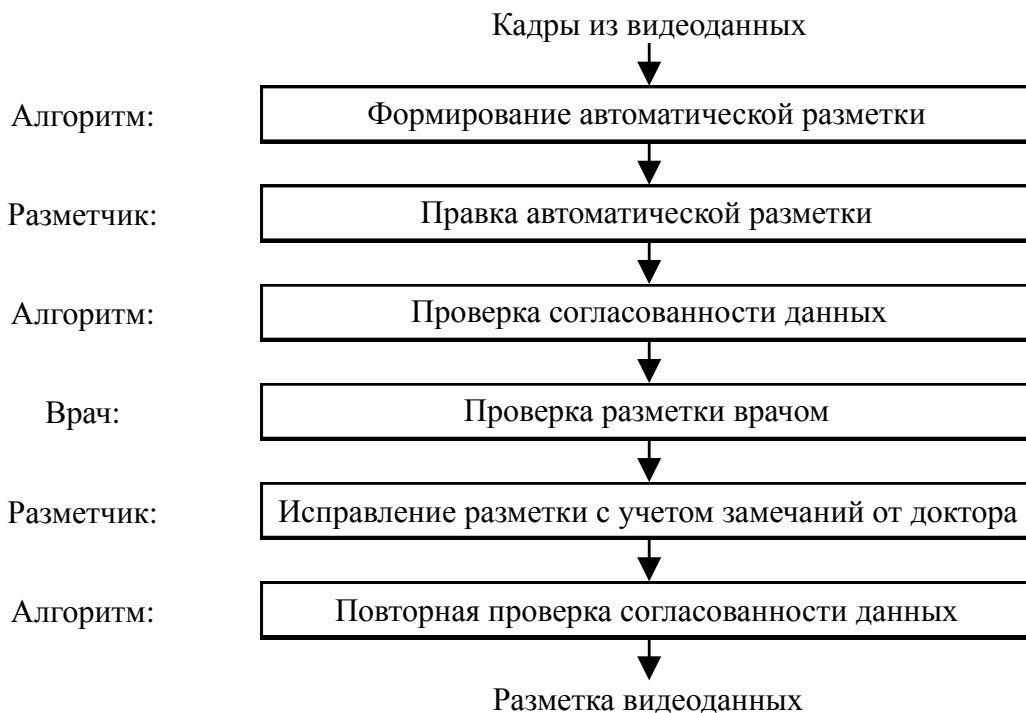


Рисунок 9 – Структурная схема методики полуавтоматической разметки областей интереса

Стоит отметить, что данную методику можно применять итеративно для разметки большого количества видеоданных. На каждой итерации стоит размечать часть данных и дообучать алгоритм на новых данных, тогда следующая итерация будет требовать меньше трудозатрат, т.к. качество автоматической разметки будет возрастать. Как показывает практика, трудозатраты при создании больших размеченных баз изображений уменьшаются при этом на 30-40%.

Используя данную методику, выполнена разметка видеофрагментов с эндоскопической системы с целью формирования тестового набора изображений из реальной клинической практики.

Для оценки качества работы итогового нейросетевого алгоритма обнаружения на базе ПРГВ использовались два различных способа. В первом из них использовались метрики средней точности для каждого обнаружения алгоритма, независимо от остальных. Во втором – для подобной оценки, как и раньше, использовались метрики точности, полноты, средней точности и  $F1$ , но с другим определением значений параметров, на основе которых они рассчитываются. Для наглядности добавим к данным метрикам и параметрам индекс  $track$ .

Определим  $TP_{track}$  как число треков из экспертной разметки, для которых есть пересечение с треками, предсказанными алгоритмом,  $FN_{track}$  – как число треков из экспертной разметки, для которых нет пересечений с треками, предсказанными

алгоритмом, а  $FP_{track}$  – как число треков, предсказанных алгоритмом, для которых нет пересечений с треками из экспертной разметки.

Таким образом, для алгоритма  $HAOO_{YXVID}$  рассчитаны метрики  $AP@0,50$ ,  $P$ ,  $R$ , и  $F1$  для баз ЯОКОБ и ПРГВ. Также для базы ПРГВ рассчитаны потрековые аналоги выбранных метрик. Значения метрик приведены в Таблице 2.

Таблица 2  
Результаты тестирования итогового алгоритма  $HAOO_{YXVID}$

Метрика	База ЯОКОБ	База ПРГВ	База ПРГВ (track)
$AP@0,50$	83,3	78,3	81,0
$P$	76,4	72,5	82,0
$R$	81,0	75,7	83,3
$F1$	78,6	74,1	82,6

Как было указано ранее, в среднем врач-эндоскопист во время диагностики обнаруживает порядка 80% полипов, что соответствует  $R_{track}=80,0\%$ . Таким образом, на основании приведенных результатов можно сделать вывод о том, что итоговый нейросетевой алгоритм демонстрирует результат обнаружения полипов, сравнимый с работой врача-эндоскописта.

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. Разработаны методика, нейросетевые алгоритмы и программное обеспечение для обнаружения патологических объектов на статических изображениях и видеоданных колоноскопических исследований в медицинской системе прикладного телевидения.
2. Предложенная модификация процесса аугментации данных позволила улучшить значение метрики  $AP@0,50$  в среднем на 1,4% для алгоритма на основе сверточной нейронной сети SSD и на 3,0% для алгоритма на основе архитектуры YOLOv5.
3. Для решения задачи обнаружения полипов на статических изображениях колоноскопических исследований предложены модификации нейросетевых алгоритмов. Их тестирование на открытой базе изображений Kvasir-SEG показало, что наилучший результат получен для алгоритма  $HAOO_{YL}$ , для которого значение метрики  $AP@[0,25..0,75]$  составило 92,5%, что превосходит известные аналоги на 5% и более. Некоторым недостатком такого алгоритма

является повышенной в 3 раза вычислительная сложность по сравнению с известным из литературы алгоритмом ColonSegNet.

4. Предложенные нейросетевые алгоритмы переобучены и протестированы на базе видеоданных колоноскопических исследований ЯОКОБ. Размер обучающего набора включал 165 665 изображений, из которых 114 754 содержат полипы, а тестовый набор – 11 192 изображения, из которых 7057 содержат полипы. Наилучший результат на тестовом наборе получен для алгоритма  $HAOO_{YX}$ , где значение метрики средней точности работы алгоритма составило 81,6%.
5. Применяя алгоритмы квантования весов СНС к алгоритму  $HAOO_{YX}$ , удалось уменьшить среднее время обработки одного кадра видеопоследовательности с 22,8 мс до 13,1 мс, что эквивалентно достижению скорости обработки в 76 кадров в секунду. При этом значение метрики  $AP@0,50$  деградировало не более чем на 0,1%. Таким образом, алгоритм  $HAOO_{YX}$  после применения методов квантования весов СНС можно использовать для обработки видеопотока с эндоскопической системы в режиме реального времени.
6. Разработан модифицированный алгоритм обнаружения полипов для работы на видеопоследовательностях  $HAOO_{YXVID}$ , позволивший улучшить значение метрики средней точности обнаружения полипов на базе видеоданных ЯОКОБ с 81,6% до 83,3% за счет применения процедуры постобработки.
7. Предложена и апробирована на наборе из более чем 31 000 видеок кадров методика полуавтоматической подготовки базы видеоданных колоноскопических исследований с разметкой областей интереса для обучения и/или тестирования нейросетевых моделей, использованная для создания базы видеоданных ПРГВ, максимально приближенной к реальным клиническим условиям. Соответствующие трудозатраты при использовании предложенной методики сокращаются на 30-40%.
8. Тестирование алгоритма  $HAOO_{YXVID}$  на базе видеоданных ПРГВ дает следующие значения метрик оценки качества:  $AP@0,50=78,3\%$ ,  $P=72,5\%$ ,  $R=75,7\%$ ,  $F1=74,1\%$  – при пок кадровом анализе работы алгоритма и  $AP@0,50_{track}=81,0\%$ ,  $P_{track}=82,0\%$ ,  $R_{track}=83,3\%$ ,  $F1_{track}=82,6\%$  – при потрековом анализе работы. Данные результаты являются практически востребованными для задачи обнаружения объектов интереса на подобном классе видеоданных.

9. Цель и задачи диссертационной работы успешно выполнены. Разработанные алгоритмы интегрированы в соответствующие программно-аппаратные комплексы. Получены два свидетельства о государственной регистрации программы для ЭВМ и свидетельство о государственной регистрации базы эндоскопических изображений.

## **ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ**

### **Статьи в изданиях, рекомендуемых ВАК**

1. Хрящев В.В., Ганин А.Н., Лебедев А.А., Степанова О.А., Кашин С.В., Куваев Р.О. Разработка и анализ алгоритма детектирования патологий на эндоскопических изображениях желудка на основе сверточной нейронной сети // Цифровая обработка сигналов. 2018. №3. С.70-75.
2. Лебедев А.А., Хрящев В.В., Кашин С.В., Среднякова А.С., Казина Е.М. Применение методов глубокого обучения для поддержки врачебного решения при эндоскопическом исследовании желудка // Модели, системы, сети в экономике, технике, природе и обществе. 2021. №2(38). С. 95-106.
3. Лебедев А.А., Хрящев В.В., Среднякова А.С., Казина Е.М. Разработка алгоритма детектирования полипов на эндоскопических изображениях с использованием сверточных нейронных сетей // Цифровая обработка сигналов. 2021. №2. С. 55-60.
4. Лебедев А.А., Приоров А.Л., Хрящев В.В., Среднякова А.С. Постобработка результатов работы нейросетевого детектора полипов на видеоизображениях колоноскопических исследований // Цифровая обработка сигналов. 2021. №4. С. 68-71.

### **Материалы конференций, индексируемых в Scopus**

5. Khryashchev V., Stepanova O., Lebedev A., Kashin S., Kuvaev R. Deep learning for gastric pathology detection in endoscopic images // ACM International Conference Proceeding Series, 3rd International Conference on Graphics and Signal Processing, ICGSP 2019. Hong Kong, 2019. P. 90-94.
6. Lebedev A., Khryashchev V., Stefanidi A., Stepanova O., Kashin S., Kuvaev R. Convolutional neural network for early detection of gastric cancer by endoscopic video analysis // Proceedings of Twelfth International Conference on Machine Vision (ICMV 2019). Amsterdam, Netherlands, 2019. P. 585-590.

7. Lebedev A., Khryashchev V., Kazina E., Zhuravleva A., Kashin S., Zavyalov D. Automatic Identification of Appendiceal Orifice on Colonoscopy Images Using Deep Neural Network // Proceedings of 2020 IEEE East-West Design & Test Symposium (EWDTS 2020). Varna, Bulgaria, 2020. P. 1-5.
8. Khryashchev V., Lebedev A., Stepanova O., Srednyakova A. Analysis of Pathologies on Endoscopic Images of the Stomach Using SSD and RetinaNet Neural Network Architecture // Proceedings of 2021 IEEE East-West Design and Test Symposium (EWDTS 2021). Batumi; Georgia, 2021. P. 119-123.

#### **Доклады на российских и международных конференциях**

9. Лебедев А.А., Степанова О.А., Юрченко Е.А., Хрящев В.В. Разработка алгоритмов анализа изображений для классификации патологий слизистой оболочки желудка // Докл. 20-й междунар. конф. и выставки «Цифровая обработка сигналов и ее применение». Москва, 2018. Т. 8. С. 223-227.
10. Степанова О.А., Лебедев А.А., Хрящев В.В., Приоров А.Л. Использование сверточной нейронной сети SSD для детектирования патологий при эндоскопии желудка // Докл. 21-й междунар. конф. и выставки «Цифровая обработка сигналов и ее применение». Москва, 2019. С. 533-537.
11. Лебедев А.А., Хрящев В.В., Казина Е.М., Среднякова А.С., Журавлева А.С. Распознавание устья червеобразного отростка на эндоскопических изображениях прямой кишки на основе сверточной нейронной сети // Докл. 22-й междунар. конф. и выставки «Цифровая обработка сигналов и ее применение». Москва, 2020. С. 638-642.
12. Лебедев А.А., Хрящев В.В., Степанова О.А. Распознавание патологий желудка по видеоизображениям эндоскопического исследования // Сб. науч. ст. по материалам XVII междунар. науч.-техн. конф. «Новые информационные технологии и системы». Пенза, 2020. С. 197-201.
13. Лебедев А.А. Разработка алгоритма детектирования полипов на видеопотоке с эндоскопической системы с использованием методов глубокого машинного обучения // Материалы 14-й междунар. науч.-техн. конф. «Перспективные технологии в средствах передачи информации». Владимир, 2021. С. 400-403.
14. Лебедев А.А., Приоров А.Л., Андержанова А.С., Среднякова А.С. Разработка алгоритмов детектирования полипов на эндоскопических изображениях // Сб.

ст. XXI междунар. науч.-техн. конф. «Проблемы информатики в образовании, управлении, экономике и технике». Пенза, 2021. С. 51-56.

15. Лебедев А.А., Казина Е.М., Журавлева А.С., Завьялов Д.В. Детектирование и классификация полипов толстой кишки на основе анализа эндоскопических изображений методами глубокого машинного обучения // Докл. 23-й междунар. конф. и выставки «Цифровая обработка сигналов и ее применение». Москва, 2021. С. 214-217.
16. Лебедев А.А. Разработка нейросетевого алгоритма обнаружения объектов на видеоизображениях в системах прикладного телевидения // Докл. 77-й всерос. конф. (с междунар. участием) «Радиоэлектронные устройства и системы для инфокоммуникационных технологий». Москва, 2022. 5 с.
17. Лебедев А.А., Хрящев В.В. Разработка алгоритма детектирования полипов на видеопотоке с эндоскопической системы в режиме реального времени // Сб. докл. 77-й науч.-техн. конф. «СПб НТО РЭС им. А.С. Попова, посвященная Дню радио». Санкт-Петербург, 2022. С. 49-52.

#### **Свидетельства о государственной регистрации**

18. Лебедев А.А., Хрящев В.В., Завьялов Д.В., Казина Е.М., Журавлева А.С. AI.ColonoscopySolution 1.0 – программа для анализа колоноскопических изображений на основе технологий машинного обучения // Свидетельство о регистрации программы для ЭВМ №2020618382 от 24.07.2020.
19. Андержанова А.С., Казина Е.М., Лебедев А.А., Среднякова А.С. Endoscopy\_Polyps 1.0 – программа для детектирования и распознавания полипов в кишечнике с использованием эндоскопических изображений // Свидетельство о государственной регистрации программы для ЭВМ №2022610262 от 11.01.2022.
20. Гусейнова С.Р., Завьялов Д.В., Лебедев А.А. Colonoscopy.DB – база данных эндоскопических изображений толстой кишки // Свидетельство о государственной регистрации базы данных №2022620355 от 28.02.2022.

Лебедев Антон Александрович

**ИССЛЕДОВАНИЕ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ  
ОБНАРУЖЕНИЯ ОБЪЕКТОВ НА ВИДЕОИЗОБРАЖЕНИЯХ  
В МЕДИЦИНСКИХ СИСТЕМАХ ПРИКЛАДНОГО  
ТЕЛЕВИДЕНИЯ**

Автореферат диссертации на соискание ученой степени  
кандидата технических наук

Подписано в печать 23.06.2022.

Формат 60x84 1/16. Усл. печ. л. 1. Тираж 100 экз.

ИП Грязнухин Р.А. 150000, Ярославль, ул. Б. Октябрьская, 37/1.

