

ОТЗЫВ

официального оппонента к.т.н., доцента Мотыко А.А. на диссертацию Лебедева Антона Александровича «Исследование нейросетевых алгоритмов обнаружения объектов на видеоизображениях в медицинских системах прикладного телевидения», представленную на соискание ученой степени кандидата технических наук по специальности 2.2.13. Радиотехника, в том числе системы и устройства телевидения

Актуальность темы диссертационного исследования

Цифровая обработка видеоизображений – одна из перспективных и быстро развивающихся областей науки и техники. Среди множества различных задач, решаемых в данной области важное значение в практических приложениях имеют задачи обнаружения и классификации объектов на видеоизображениях. Такие алгоритмы имеют широкое применение в различных прикладных телевизионных системах (ПТС) – от систем охранного телевидения, до модулей поддержки принятия решения в медицинских видеокомпьютерных системах. Для решения задач обнаружения объектов широко используются алгоритмы машинного обучения, среди которых в последние годы доминирующее положение заняли алгоритмы на основе сверточных нейронных сетей. Разработка и исследование таких алгоритмов, анализирующих видеопоток в режиме реального времени, представляет собой актуальную научно-техническую задачу в области построения современных ПТС.

Медицинские ПТС занимают важную роль в эндоскопических исследованиях желудочно-кишечного тракта. Разработка нейросетевых алгоритмов в данной области осложняется отсутствием или недостатком открытых баз видеопоследовательностей с экспертной разметкой областей интереса в каждом кадре. Получение таких баз данных трудоемкий процесс, требующий больших трудозатрат не только со стороны специалистов по работе с изображениями, но и со стороны врачей-эндоскопистов.

Таким образом, тема диссертации для сегодняшнего уровня развития систем прикладного телевидения в медицине является актуальной. Ее развитие в

диссертации Лебедева А.А. имеет важное теоретическое и практическое значение для ряда современных научно-технических областей.

Структура 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.

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 проведен обзор существующих методов и алгоритмов обнаружения объектов на изображениях, а также рассмотрены критерии качества, используемые для подобных алгоритмов. Рассмотрены как классические алгоритмы машинного обучения, так и методы глубокого обучения (на основе применения сверточных нейронных сетей). Описаны современные ПТС и их применение в медицине. Сформулирована задача обнаружения полипов кишечника на изображениях, полученных в эндоскопической ПТС.

Во второй главе предложены алгоритмы обнаружения полипов на эндоскопических изображениях на основе архитектур сверточных нейронных сетей SSD и YOLOv5. Приведены результаты их обучения и тестирования на общедоступном наборе эндоскопических изображений Kvasir-SEG. Предложена модификация алгоритма аугментации данных с учетом специфики эндоскопических изображений. Полученные результаты показывают, что наилучший результат получен для предложенного нейросетевого алгоритма $HAOO_{YL}$, для которого значение метрики $AP@[0,25..0,75]$ составляет 92,5%, что превосходит известные аналоги на 5% и более.

В третьей главе рассматривается алгоритм постобработки результатов обнаружения полипов на эндоскопических видеопоследовательностях, позволяющий выполнять связывание обнаружений между последовательными кадрами в треки. Данный подход позволяет повысить точность работы алгоритма при анализе видеопоследовательностей. Представлены результаты тестирования

предложенного алгоритма обнаружения полипов на двух собранных автором базах эндоскопических видеоизображений ЯОКОБ и ПРГВ. Показано, что точность работы предложенного алгоритма обнаружения полипов на видеоизображениях эндоскопических исследований является сравнимым с аналогичной экспертной работой врача-эндоскописта.

В заключении приведены основные результаты, полученные в диссертационной работе.

Наиболее значимые **научные результаты диссертационной работы** состоят в следующем:

1. Предложенный нейросетевой алгоритм обнаружения полипов превосходит существующие аналоги на открытом наборе эндоскопических изображений Kvasir-SEG.
2. Предложен алгоритм аугментации и постобработки видеоизображений, учитывающий специфику видеоэндоскопических данных.
3. Предложена методика подготовки базы видеоданных эндоскопических исследований с разметкой областей интереса для обучения и/или тестирования нейросетевых моделей, позволяющая сократить соответствующие трудозатраты на 30-40%.

Достоверность полученных научных результатов обеспечивается корректным использованием математического аппарата, полученными результатами компьютерного моделирования и их сопоставления с известными из научно-технической литературы данными.

Практическая значимость результатов

Практическая значимость полученных результатов состоит в следующем:

1. Предложенный нейросетевой алгоритм обнаружения полипов демонстрирует высокие значения стандартных метрик оценки качества на базе видеоданных, максимально приближенных к реальным клиническим условиям.
2. Разработанный алгоритм позволяет проводить анализ видеопотока с эндоскопической системы в режиме реального времени.

3. Собраны и экспертно размечены базы видеоизображений ЯОКОБ и ПРГВ, которые могут в дальнейшем использоваться для обучения и/или тестирования нейросетевых моделей.

Замечания по диссертационной работе

1. Из текста работы неясно, почему с практической точки зрения необходимо решать задачу не только обнаружения, но и точной локализации полипа в кадре видеопоследовательности.
2. На с. 31 описывается, что общепринятой метрикой качества для алгоритмов обнаружения объектов является метрика $mAP@[0,5..0,95]$, усреднённая по ряду значений IoU от 0,5 до 0,95 с шагом 0,05, однако в дальнейшем в работе используется метрика $mAP@[0,25..0,75]$, усреднённая по ряду значений IoU от 0,25 до 0,75 с шагом 0,05.
3. В третьей главе не приводится оценка качества работы алгоритма постобработки, т.е. не показано насколько корректно происходит связывание обнаружений в треки.
4. На с. 77 не дано объяснений, как была сформирована обучающая и тестовая выборка. Не ясно, отличалась ли тестовая выборка до и после дообучения на базе ЯОКОБ. Не было ли пересечений видеофрагментов в тренировочной и обучающей выборках.
5. На с. 85 для предложенного алгоритма с постобработкой не объяснено, как планируется его реализация для режима реального времени, непосредственно при проведении эндоскопического обследования.

Указанные недостатки в целом не влияют на оценку диссертации как законченной научно-квалификационной работы, содержащей решение важной научной задачи.

Заключение

Автореферат отражает основное содержание диссертации.

Результаты работы в достаточном объеме опубликованы и апробированы на научных и научно-технических конференциях различного уровня.

Диссертационная работа «Исследование нейросетевых алгоритмов обнаружения объектов на видеоизображениях в медицинских системах прикладного телевидения» соответствует требованиям ВАК, предъявляемым к диссертациям на соискание ученой степени кандидата технических наук, а ее автор Лебедев Антон Александрович заслуживает присуждения ученой степени кандидата технических наук по специальности 2.2.13. Радиотехника, в том числе системы и устройства телевидения.

Официальный оппонент

кандидат технических наук, доцент,

доцент кафедры телевидения

и видеотехники СПб ГЭТУ «ЛЭТИ» Александр Александрович Мотыко



Дата: 18.08.2022

Адрес: 199106, г. Санкт-Петербург, ул. Гаванская, д. 18, к. 95

Телефон: +7(905)-228-90-82

e-mail: motyko.alexandr@yandex.ru



Подпись доцента А.А. Мотыко заверяю:

Начальник отдела кадров

СПб ГЭТУ «ЛЭТИ»


