

ОПЫТ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОТРАБОТКЕ И ОЦЕНКЕ БАЗОВЫХ НАВЫКОВ ЭНДОХИРУРГА

Горшков Максим Дмитриевич¹, Ким Евгений Владимирович², Бесчастнов Игорь Владимирович³,
Тюрин Ростислав Русланович⁴, Павлов Сергей Алексеевич⁵

1. Российское общество симуляционного обучения в медицине РОСОМЕД, г. Москва, Российская Федерация; Европейский институт симуляции в медицине, г. Штутгарт, Германия
2. Учебно-симуляционный центр Синтомед, г. Москва, Российская Федерация
3. НИИ Механики, МГУ им. М. В. Ломоносова, г. Москва, Российская Федерация
4. Механико-математический факультет МГУ им. М. В. Ломоносова, г. Москва, Российская Федерация
5. Факультет космических исследований МГУ им. М. В. Ломоносова, г. Москва, Российская Федерация

ORCID: Горшков М. Д. — 0000-0003-0446-0787

ORCID: Ким Е. В. — 0000-0001-6689-2310

ORCID: Бесчастнов И. В. — 0009-0008-8847-9314

ORCID: Тюрин Р. Р. — 0009-0007-9137-5340

ORCID: Павлов С. А. — 0009-0002-8983-4564

gorshkov@rosomed.ru

DOI: 10.46594/2687-0037_2025_2_2006

Аннотация. В статье рассматривается применение технологий искусственного интеллекта для отработки и объективной оценки базовых навыков эндохирурга на симуляционной платформе БЭСТА-2.0. Представлены архитектура системы, алгоритмы технического зрения, этапы обучения нейросетей и результаты их применения. В ходе трёхэтапного исследования были достигнуты высокие показатели точности автоматической оценки (до 95% совпадения с экспертной), а также продемонстрирована высокая воспроизводимость результатов. Обсуждаются ключевые ограничения и пути их устранения. Автоматизация оценки эндохирургических навыков позволяет повысить доступность качественного обучения, разгрузить преподавателей и обеспечить немедленную обратную связь в масштабируемом формате. Работа подчёркивает перспективность дальнейшего внедрения ИИ в лапароскопический тренинг и подготовку хирургов в целом.

Ключевые слова: БЭСТА-2.0, ИИ, автоматическая объективная оценка, виртуально-дополненная реальность, искусственный интеллект, нейросети, машинное обучение, компьютерное зрение, техническое зрение, медицинский тренажёр, лапароскопическая хирургия, симуляционное обучение, хирургическое обучение, лапароскопический тренинг.

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Для цитирования: Горшков М. Д. Опыт применения искусственного интеллекта в отработке и оценке базовых навыков эндохирурга // Виртуальные технологии в медицине. 2025. № 2. DOI: 10.46594/2687-0037_2025_2_2006

Научные специальности: 3.1.9. Хирургия; 3.2.3. Общественное здоровье и организация здравоохранения, социология и история медицины.

Поступила в редакцию 02 апреля 2025 г.

Поступила после рецензирования 10 июня 2025 г.

Принята к публикации 16 июня 2025 г.

APPLICATION OF ARTIFICIAL INTELLIGENCE IN TRAINING AND ASSESSMENT OF BASIC ENDOSCOPIC SURGICAL SKILLS

Gorshkov Maxim¹, Kim Evgeny², Beshchastnov Igor³, Tyurin Rostislav⁴, Pavlov Sergey⁵

1. Russian Society for Simulation Education in Medicine (ROSOMED), Moscow, Russian Federation; European Institute for Simulation in Medicine (EuroMedSim), Stuttgart, Federal Republic of Germany
2. Sintomed Training and Simulation Center, Moscow, Russian Federation
3. Research Institute of Mechanics, Lomonosov Moscow State University, Moscow, Russian Federation
4. Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Moscow, Russian Federation
5. Faculty of Space Research, Lomonosov Moscow State University, Moscow, Russian Federation

gorshkov@rosomed.ru

DOI: 10.46594/2687-0037_2025_2_2006

Annotation. This article examines the application of artificial intelligence technologies for training and objectively assessing basic laparoscopic skills using the BÉSTA 2.0 simulation platform. The paper presents the system architecture, computer vision algorithms, stages of neural network training, and performance results. A three-phase study demonstrated high assessment accuracy (up to 95% agreement with expert ratings) and strong reproducibility of outcomes. Key limitations and mitigation

strategies are discussed. Automated evaluation of the endosurgical skills enables scalable, instructor-independent feedback, improves training efficiency, and ensures real-time insight into trainee performance. The findings emphasize the potential of incorporating AI into laparoscopic training and, more broadly, into contemporary surgical education.

Keywords: BESTA, AI, AR, automatic objective assessment, augmented reality, virtually enhanced reality, artificial intelligence, neural networks, machine learning, computer vision algorithms, technical vision, medical simulator, laparoscopic surgery, simulation training, surgical education, laparoscopic training.

Conflict of interest. The author declares no conflict of interest.

For quotation: Gorshkov M. Application of Artificial Intelligence in Training and Assessment of Basic Endoscopic Surgical Skills // Virtual Technologies in Medicine. 2025. No. 2. DOI: 10.46594/2687-0037_2025_2_2006

Received April 02, 2025

Revised June 10, 2025

Accepted June 16, 2025

Введение

Отработка мануальных эндохирургических навыков более 20 лет осуществляется с использованием коробочных видеотренажёров и виртуальных симуляторов [3, 6, 7]. Обучение этим навыкам представляет собой трудоёмкий и методически сложный процесс, начинающийся с освоения базовых манипуляций. В целях стандартизации и структурирования подготовки специалистов Российское общество симуляционного обучения в медицине (РОСОМЕД) совместно с Российским обществом хирургов и Российским обществом эндохирургов в 2015 г. разработало курс «Базовый эндохирургический симуляционный тренинг и аттестация» БЭСТА-2.0, представленный в 2016 г. на XIX съезде Российского общества эндохирургов [1].

В процессе многократного повторения упражнений необходим постоянный контроль качества их выполнения и своевременное внесение коррективов в действия обучающихся. При занятиях на коробочных тренажёрах эту функцию выполняет эксперт-наставник, осуществляющий методическое сопровождение и при необходимости корректирующий технику выполнения. Виртуальные симуляторы, в свою очередь, предоставляют автоматическую обратную связь, однако их высокая стоимость ограничивает широкое внедрение и доступность подобных решений.

В последние годы технологии искусственного интеллекта (далее — ИИ) активно внедряются в медицинское образование, включая хирургическое обучение. Одним из перспективных направлений является автоматизация оценки лапароскопических навыков с применением компьютерного зрения [4]. Автоматизированные системы, основанные на алгоритмах ИИ и работающие в режиме реального времени, анализируют действия обучающихся, сравнивая их с эталонными образцами, и на основе заданных параметров выставляют оценку, что позволяет снизить нагрузку на преподавателей, стандартизировать процесс обучения и в конечном счете повысить его эффективность.

В статье рассматривается опыт разработки и применения системы машинного зрения (также называемой техническим или компьютерным зрением), предна-

значенной для анализа действий обучаемого (далее — Система). С целью проведения объективной оценки выполнения упражнений по отработке базовых эндохирургических навыков Система встроена в тренажёр виртуально-дополненной реальности БЭСТА-2.0. Комплекс состоит из аппаратной части (бокс-тренажёр с двумя камерами, комплект лапароскопических инструментов и оригинальных учебных пособий) и программного обеспечения, одним из компонентов которого является модуль распознавания и оценки действий обучаемого на основе ИИ. Ниже описаны архитектура и функциональность Системы, методика сбора и разметки данных, алгоритмы обучения нейросетей, а также приведены результаты её валидации на выборке более 300 видеозаписей учебных упражнений.

Цель

Цель исследования — разработать и оценить эффективность Системы автоматизированной отработки и объективной оценки выполнения упражнений курса БЭСТА-2.0. Система должна с помощью технического зрения (ИИ) распознавать ключевые объекты, действия и события в реальном времени и по завершении упражнения фиксировать допущенные ошибки и формировать количественную оценку уровня его выполнения, сопоставимую с экспертной. Необходимо добиться высокой степени корреляции (не менее 90%) между автоматической оценкой и заключением экспертов, а также обеспечить универсальность системы для различных типов упражнений (навигация лапароскопом, перемещение объектов, иссечение круга, наложение клипс и эндоскопической петли).

Материалы и методы

Учебно-тренировочная система БЭСТА-2.0 (рис. 1) состоит из эндовидеохирургического бокса-тренажёра с видеокамерами, компьютера, монитора и набора инструментов с учебными пособиями для выполнения комплекса из 10 упражнений согласно заданному алгоритму БЭСТА-2.0 (см. табл. 1). Подробно методики выполнения, возможных ошибок, критериев оценки и проходных баллов по каждому из 10 упражнений изложены в книге общества РОСОМЕД [2].

С апреля 2023 года в конструкцию бокса было внесено ряд изменений для улучшения дидактического функ-

Упражнения программы БЭСТА-2.0

Задание	Описание учебного пособия и упражнения
1. Навигация лапароскопом	Пирамидка с 9 ячейками, имеющими два набора цифр. Первый набор — нумерация, идущая по порядку, а второй набор цифр идёт вразнобой и скрыт от прямого обзора. Необходимо пройти от одной ячейки к другой согласно последовательности, заданной скрытыми цифрами, которую можно определить только с помощью осмотра скошенной оптикой
2. Перенос по штырькам	Платформа с 12 штырьками и 6 призмами. Необходимо снять призму со штырька, перехватить другим инструментом и, не уронив её, надеть на штырь с противоположной половины. После переноса 6 призм в одну сторону действия выполняются в обратной последовательности
3. Бимануальная координация	Пирамидка из упражнения 1 и набор 9 крышек с эластичными рукоятками для захвата браншами. Диссектором открываются крышки, а лапароскопом определяется скрытый номер. Далее выполнение как в упражнении 1
4. Иссечение круга	Салфетка с двумя метками-окружностями, между которыми с помощью эндоножниц необходимо прорезать круг, не затрагивая меток
5. Клипирование и пересечение	Платформа как в упражнении 2, на штырьках которой растянуты две резинки. Необходимо наложить клипсы и между ними пересечь структуры
6. Захват и прошивание	Платформа как в упражнении 2, на штырьках которой растянуты две резиновые ленты, которые необходимо прошить в шести точках
7. Экстракорпоральный шов	Прошивание модели кишки по заданным меткам и затягивание толкателем узлов, сформированных экстракорпорально. Кишку имитирует дренаж Пенроуза с отверстием и метками
8. Эндопетля	Поролоновая форма с тремя отростками, центральный из которых имеет маркировку. Необходимо накинуть поверх отростка эндоскопическую петлю и затянуть её точно поверх метки
9. Интракорпоральный шов	Прошивание модели кишки по заданным меткам и интракорпоральное формирование и затягивание узлов. Кишку имитирует дренаж Пенроуза с отверстием и метками
10. Непрерывный (кисетный) шов	Прошивание модели кишки по заданным меткам и интракорпоральное формирование непрерывного (кисетного) шва. Кишку имитирует дренаж Пенроуза с отверстием и метками



Рис. 1. Выполнение упражнения на системе БЭСТА-2.0

ционала Системы и модернизации аппаратной части для целей технического зрения. Так, в бокс встроена стереоскопическая камера, способная определять положение объектов в трёхмерном пространстве. Стереоскопическая камера обеспечивает более точный трекинг движений инструментов по трём осям и фиксацию ряда параметров (траектория перемещения, направление движения инструмента и др.), что существенно важно для точности объективной оценки, дальнейшего анализа выполненных действий и распознавания ошибок на основе ИИ. Кроме того, незначительно изменены габариты бокса, добавлена разметка, а количество рабочих отверстий увеличено с 2 до 6, что позволило отрабатывать манипуляции с инструментами, введёнными под разными углами и в разных положениях портов, а также имитировать командную работу с ассистентом или оператором камеры. Для заданий по отработке навыков владения скошенной оптикой используется медицинский лапароскоп 10 мм / 30 град., на окуляре которого с помощью адаптера (coupler) фиксируется видеокамера. Таким образом, в тренажёре версии БЭСТА-2.0 используется три видеокамеры — общая для 8 упражнений из 10, камера лапароскопа — для двух заданий с использованием эндоскопической оптики и стереоскопическая, предоставляющая данные для анализа техническим зрением.

Архитектура программных модулей. Программная часть представлена предобученными моделями с блоком исполнения (оценки), реализованным на языке Python. В архитектуру входят несколько моделей компьютерного зрения, каждая из которых отвечает за распознавание или измерение определённого параметра упражнения. Результаты работы моделей агрегируются и интерпретируются модулем оценки, который выдаёт итоговые баллы и отчёты об ошибках. Общая цепочка обработки данных выглядела следующим образом: видео с выполнением упражнений записывались и собирались на компьютере тренажёра, после чего загружались в облачное хранилище (формат S3). Затем данные передавались экспертам-разметчикам: при обучении детекторов и классификаторов объектов — с покадровой разметкой, а при обучении модулей, оценивающих результаты упражнения в целом, — с разметкой, исходя из всего видео и итогового результата. После разметки все полученные изображения и метаданные (координаты объектов, классификация состояний, результаты упражнений и пр.) сохранялись в финальном облачном хранилище, откуда автоматически загружались для обучения нейросетей.

Сбор данных и их разметка. Для обучения системы в три этапа были собраны данные выполнения упражнений. На **первом этапе** опытные лапароскопические хирурги (N = 5) выполнили эталонные примеры упражнений и распространённые ошибочные действия, всего записано около 50 видео. Эти видеозаписи послужили исходным обучающим набором для черновой версии алгоритмов. Разметка включала: количество допущенных ошибок (например, число падений объектов), положения важных объектов в кадре (инструменты, учебные детали) и принадлеж-

ность действий к правильным или ошибочным. Была проделана продолжительная и масштабная работа по набору материала для машинного обучения — видеозаписей выполнения упражнений. Задания выполнялись студентами медицинских вузов из хирургических научных кружков и просто увлекающимися хирургией, а также ординаторы и курсанты, которые обучались по программе освоения базовых лапароскопических навыков. На **втором этапе** к сбору данных привлекались обучающиеся хирурги с базовыми навыками (N = 24), выполнившие по пять полных попыток каждого из выбранных упражнений. Таким образом, общий объём данных после второго этапа составил более 170 полноценных видеоопераций. Наконец, на **третьем этапе** были привлечены студенты и ординаторы, ранее не имевшие опыта лапароскопических вмешательств (N = 27), также выполнившие по пять попыток упражнений. Суммарно набор данных для обучения и валидации системы значительно превысил 300 видеозаписей 56 пользователей. Дополнительно в систему загружались многочисленные фрагменты видео (неполные выполнения упражнений, количество которых не подсчитывалось) для расширения вариативности траекторных данных, однако при расчётах точности они не учитывались. В исследовании приняли участие Симуляционный центр «Синтомед», факультет фундаментальной медицины Московского государственного университета, Сеченовский университет, Российский университет медицины, Российский национальный исследовательский медицинский университет — г. Москва; Приволжский исследовательский медицинский университет, г. Нижний Новгород, и ряд других образовательных организаций.

Каждое видео было размечено для обучения соответствующих моделей. Разметка осуществлялась вручную специалистами: для моделей детекции и сегментации — разметка каждого кадра (контуры инструментов, объекты, маркеры), для моделей классификации — определение класса состояния по ключевым кадрам, для алгоритмов оценки — подсчёт целевых метрик по видео, например реальное число допущенных ошибок или измерение вручную длины траектории, точности и площади иссечения круга. Такая трудоёмкая разметка оправдана необходимостью предоставить алгоритмам достаточный объём данных для обучения на уровне, не уступающем человеческим оценкам.

Модели и алгоритмы ИИ. В разработанной системе использован комплекс разнородных алгоритмов: как классического компьютерного зрения, так и методов машинного обучения. Для относительно простых задач оценки применялись алгоритмы без глубокого обучения. Например, упражнение «Работа с камерой» (навигация камерой) обрабатывалось с помощью нейросетей, а упражнение «Эндоскопическая петля» (наложение эндоскопической петли типа Рёдера и затягивание узла на заданной отметке) — наоборот, с помощью традиционных методов компьютерного зрения: определялось положение наложенной петли относительно нанесённой на объект метки и вычислялось отклонение от целевой точки в мм. В отличие от нейросетей, классические алгоритмы требуют значи-

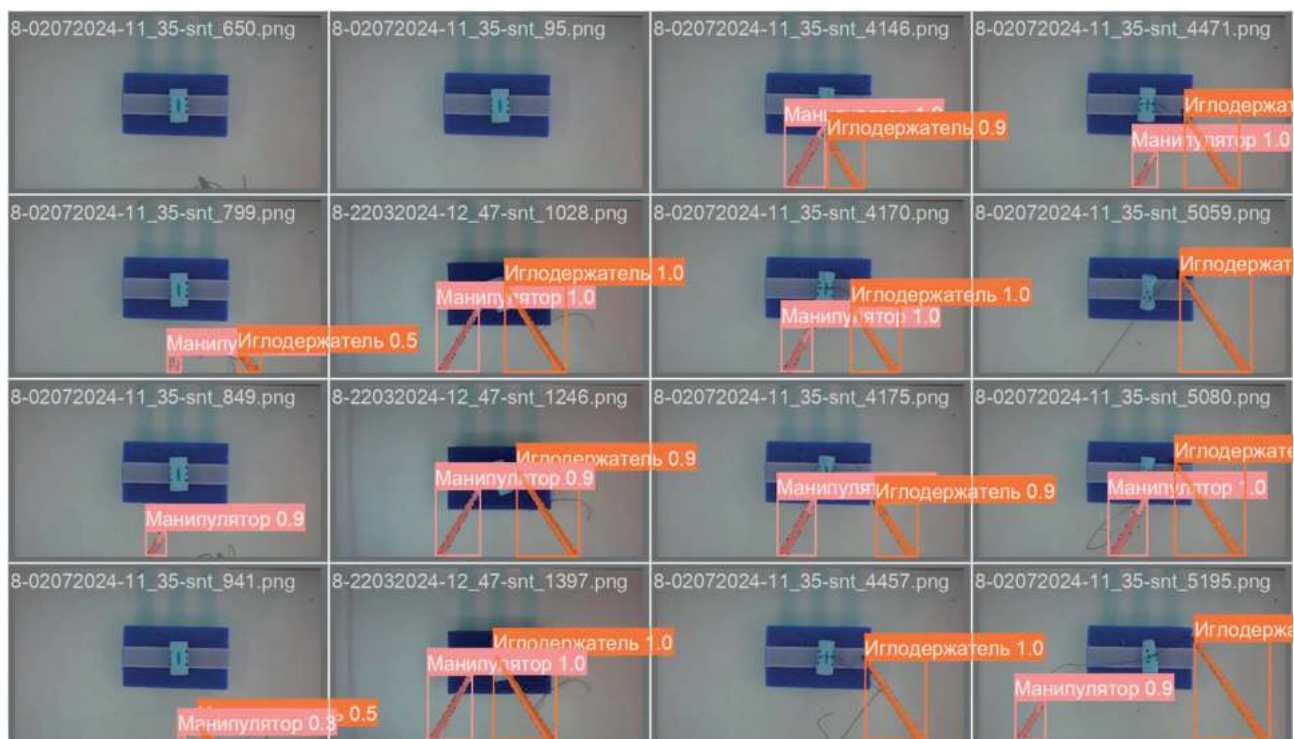


Рис. 2. Распознавание инструментов компьютерным зрением

тельно меньшего количества данных и вычислительных мощностей для обучения, легче интерпретируются и в основном работают с геометрией и цветами на изображении.

В более сложных упражнениях и метриках, где простые алгоритмы не обеспечивали нужной точности, были внедрены обучаемые модели. Были разработаны, в частности, следующие модули на основе ИИ:

- **Сегментация инструментов.** Задача состояла в выделении в видеопотоке лапароскопических инструментов — манипуляторы сегментировались целиком как один объект, и далее считалось, что манипулятор всегда правильно ориентирован рабочей частью внутрь. Обучающая выборка составила ~50 000 аннотированных кадров, содержащих инструменты в разных положениях. Модель выполняет семантическую сегментацию изображения на пять классов. В результате на каждом кадре выделяются области, соответствующие правому и левому инструменту, что позволяет точно измерять положения концов инструментов и траектории их движения (рис. 2).
- **Распознавание символов.** Для двух упражнений по отработке владения лапароскопом со скошенной под 30 градусов оптикой требовалось распознавать метки или цифры, видимые через эндоскопическую камеру. Была обучена модель для детекции цифр в кадре (номеров ячеек и скрытых от прямого обзора цифр). Она представляет собой классификатор на основе свёрточной нейросети, распознающий 10 классов цифр. Это позволило автоматически фиксировать выбор определённых позиций или мишеней, обозначенных цифрами, во время упражнений.

- **Детекция объектов (призм).** В упражнении «Перемещение объектов» используются шесть разноцветных силиконовых призм с отверстиями, которые нужно переместить с шести штырьков с одной стороны подставки на другую, а затем в обратном порядке, передавая их на весу из бранш одного инструмента в бранш другого. Обучение выполнялось на ~40 000 кадров с разметкой этих состояний.
- Алгоритм подсчёта падений оказался чувствителен к настройке гиперпараметров (масштаб времени, допустимое расстояние для кластера и др.), поэтому для его калибровки использовался перебор параметров по аналогии с GridSearch. За метрику качества при подборе параметров бралась средняя ошибка в количестве падений по всем видео или доля видео с неверным подсчётом. Оптимизированный таким образом алгоритм достигает совпадения с оценкой экспертов приблизительно в 90% случаев.
- **Классификация положения клипсы.** Упражнение «Клипирование и пересечение» имитирует наложение клипс на «сосуд» с последующим его пересечением эндоножницами между ними. В тренажёре данная манипуляция реализована с помощью двух эластичных резинок, растянутых между штырьками, на которые накладываются стандартные эндохирургические клипсы размера ML. Различная форма фигуры, образованной двумя резинками после каждого этапа (нет клипс, наложена первая клипса, вторая клипса, пересечение выполнено и т. д.), определяет этап и успешно выполняется. Для данного упражнения разработан классификатор состояний резинок по изображению. Выделено шесть классов,

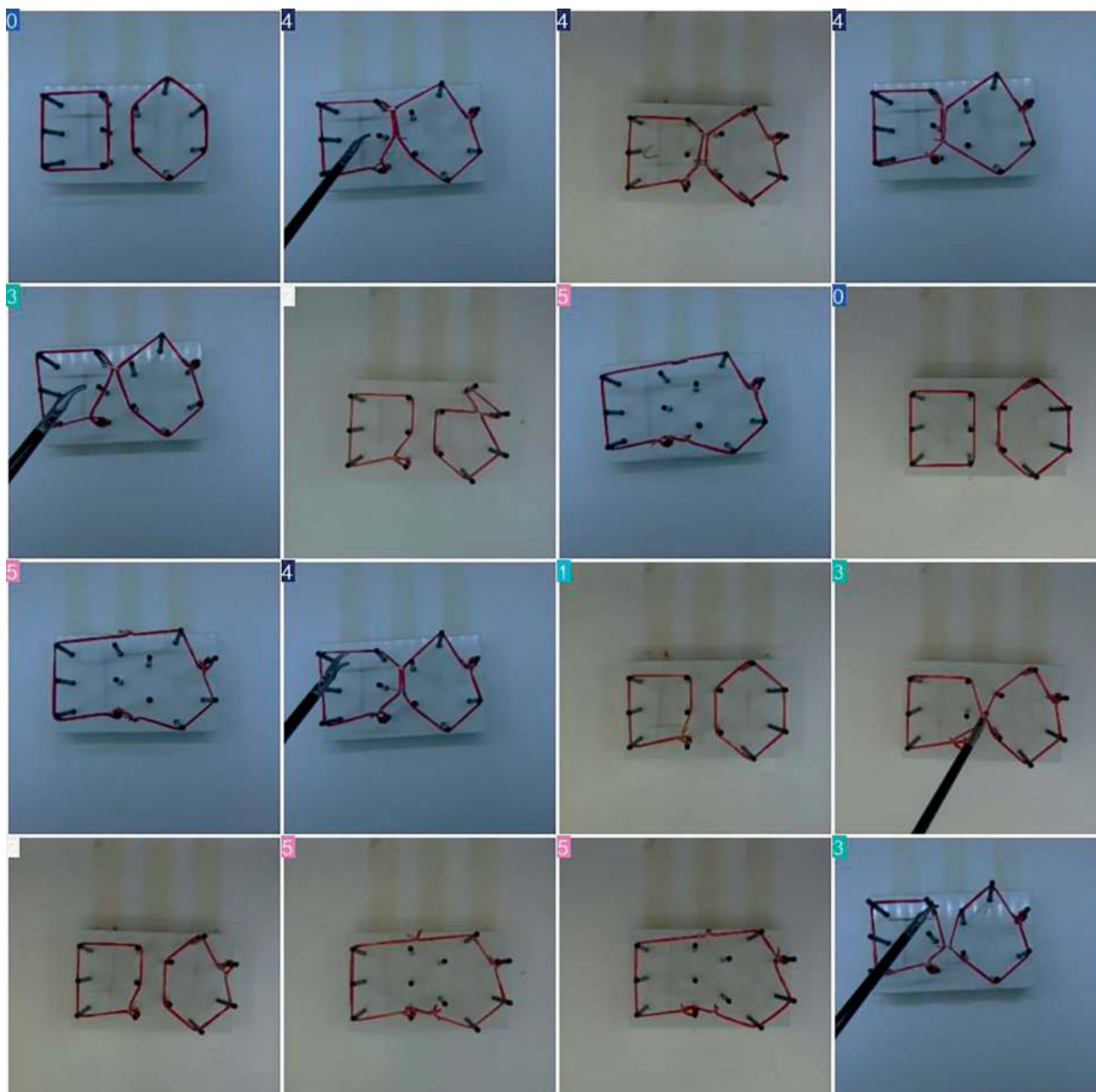


Рис. 3. Определение этапа выполнения упражнения по классификации геометрии фигуры

отражающих этапы: от исходного состояния до правильно наложенных двух клипс и завершённого пересечения (рис. 3). Обучающая выборка относительно невелика (около 200 кадров), поэтому применялись методы увеличения данных и дообучение модели на похожих задачах, чтобы избежать переобучения. Модель показала способность различать правильное и неправильное наложение клипс и факт пересечения с достаточной для практики точностью.

- **Точность иссечения круга.** В упражнении «Иссечение круга» используется нетканая салфетка с напечатанными на ней двумя чёрными окружностями, изогнутые эндоножницы Метценбаума и эндозажим. С помощью инструментов необходимо вырезать круг точно между двумя метками,

не задевая и не заходя за них. Для определения точности иссечения реализован алгоритм обработки изображения, основанный на методах кластеризации DBSCAN и статистического анализа формы объекта. Угловое пространство делится на равные сектора, и для каждого из них рассчитывается плотность — количество точек, попавших в данный сегмент. Затем определяется средняя плотность и её вариабельность (разброс значений по секторам). Сектора, в которых плотность существенно ниже среднего, считаются участками разрыва. Разрывы объединяются в последовательности, ограниченные с обеих сторон непрерывными участками. По результатам анализа вычисляется общее число таких разрывов и их доля по отношению к общему числу сегментов. Итоговая оценка

в процентах отражает степень целостности вырезанной окружности и снижается пропорционально числу обнаруженных дефектов.

Валидация и метрики. Качество работы всех алгоритмов оценивалось на отложенной части набора видео с привлечением «золотого стандарта» — оценок экспертов по каждой попытке. Проверка точности прошла итеративно параллельно с разработкой: по мере создания нового алгоритма для упражнения он запускался на всех имеющихся размеченных видео, а результаты автоматически сравнивались с эталонными значениями метрик. Основными метриками качества служили процент совпадения автоматической оценки с экспертной и средняя величина ошибки (например, разницы в количестве ошибок или отклонения в миллиметрах). Эталонная оценка ошибок выполнялась экспертами «вручную» на основе видеозаписей упражнений. Определение реальной длины траектории инструментов и положения в пространстве выполнялось с помощью отдельных тестов по эталонному объекту — многократные повторяющиеся движения инструментов в пределах прямоугольной площадки со сторонами 100 x 60 мм определялись автоматически и сравнивались с расчётными. Если алгоритм выдавал большой разброс ошибок, вводились скорректированные метрики. Например доля видео без ошибок оценки, чтобы отдать предпочтение настройкам, обеспечивающим правильный результат в большинстве случаев, даже ценой редких значительных отклонений. Итоговая интегральная оценка формировалась Системой как совокупность точностных показателей по всем упражнениям.

Результаты и обсуждение

Точность автоматической оценки. Разработанная система технического зрения продемонстрировала на финальном этапе высокую точность автоматической оценки и её совпадение со значениями, определёнными экспертами. В среднем по всем упражнениям автоматические оценки совпадали или незначительно отличались от оценок преподавателей в более 90% случаев. Однако данная точность потребовала трёхэтапного подхода и трёхэтапного подхода и тщательной работы с данными и алгоритмами.

После первого этапа обучения (около 50 видео) алгоритмы ещё не были достаточно надёжны. Так, Система фиксировала 14 падений объектов вместо двух фактически допущенных, величина ошибки в определении положения эндопетли достигала 3 см, а при клипировании обнаружился программный сбой — программа иногда самопроизвольно завершала анализ до окончания выполнения упражнения. Сбой был устранён, а с ростом объёмов данных и оптимизацией параметров точность существенно улучшилась, например после второго этапа обучения на ~170 видео погрешность в подсчёте числа падений объектов существенно снизилась, а после третьего этапа в подавляющем большинстве подходов падения определялись верно. По мере обучения предельная точность определения отклонения положения узла петли от целевой метки сократилось сначала до 3 мм и в итоге достигло 0,7 мм. Ал-

горитм анализа клипирования со второго этапа также стал стабильно работать без ошибок, а на третьем этапе (общее число видео > 300) итоговое состояние резинки стало фиксироваться корректно во всех случаях.

Система успешно прошла тестирование на следующих базовых упражнениях эндхирургического курса БЭ-СТА-2.0: «Работа с камерой», «Перемещение призм», «Бимануальная координация», «Исечение круга», «Наложение эндоскопической петли», «Клипирование и пересечение». Помимо специфических метрик для этих шести упражнений, Система собирает и другие показатели, характеризующие технику выполнения. Во всех 10 упражнениях, а также при выполнении любой другой произвольной манипуляции Системой выполняется автоматическое определение базовых параметров: «Длительность выполнения упражнения», «Длина траектории движения правого и левого инструментов», «Соотношение этих траекторий друг с другом», «Характерная скорость».

Длительность выполнения упражнения определялась как количество секунд от начала активности на экране (введён первый инструмент, старт) и до завершения любых действий в боксе (финиш). Был выбран допустимый период отсутствия активности длительностью в 10 сек., чтобы предотвратить ложное срабатывание и преждевременное завершение анализа.

Автоматическое определение длины траектории инструментов и их положения в пространстве проверялось с помощью эталонного объекта — многократными повторяющимися движениями инструментов в пределах прямоугольной площадки со сторонами 100 x 60 мм. Отклонение составило 2,1 мм, что с учётом размеров инструмента мы посчитали допустимой погрешностью. Стереоскопическая камера позволила измерять траекторию инструментов в 3D, благодаря чему удалось оценить длину пути инструмента и в вертикальной плоскости.

Попытка измерить среднюю и максимальную скорость движения инструмента натолкнулась на факт существования артефактов. Так, при введении инструмента в бокс и его извлечении скорость резко возрастает. Простое арифметическое вычисление средней и максимальной скорости даёт неестественно высокие показатели. В связи с этим мы ввели понятие *характерной скорости движения инструментов* — расчётную метрику, определяющую сглаженную от резких движений и рывков инструментов величину, более точно передающую истинную скорость движений в ходе основных манипуляций.

Также мы определяем процентное соотношение длины траекторий инструментов в правой и левой руках. Данный показатель, показывающий, насколько активно используются обе руки, косвенно свидетельствует об *амбидекстрии* пользователя — способности одинаково хорошо владеть обеими руками. У начинающих хирургов этот показатель может сильно различаться (одна рука доминирует), тогда как у опытных приближается к 50 : 50 — обе руки работают слаженно.



Рис. 4. Пирамидка из девяти ячеек с двумя наборами цифр и набор крышек

Навигация скошенной оптикой. В данном упражнении используется пособие оригинальной разработки — пирамидка из девяти ячеек, расположенных в форме треугольника (рис. 4). Каждая ячейка содержит по две цифры, т. е. всего 18 цифр, двумя наборами по 9. Первый набор — нумерация, идущая подряд, — хорошо видна, тогда как второй набор цифр идет вразнобой и скрыт от прямого обзора. Необходимо пройти от одной ячейки к другой согласно последовательности, заданной скрытыми цифрами, которую можно определить только с помощью осмотра скошенной оптикой, меняя положение лапароскопа и вращая им.

Чтобы реалистично имитировать навигацию оптикой в ходе вмешательства пришлось отказаться от более простой и экономичной версии оптики с расположением видеокамеры на конце, в области объектива (chip-on-tip) и использовать более дорогой вариант оптической системы «лапароскоп — соединительный адаптер — видеокамера». Эта конструкция позволяет максимально реалистично воспроизводить действия ассистента на камере в ходе лапароскопического вмешательства — при вращении лапароскопа для осмотра недоступных участков зоны вмешательства по-

зволяет сохранять положение камеры и «удерживать горизонт» поля зрения на экране монитора.

На первых этапах Система не показала удовлетворительной точности. Связано это было с тем, что пользователь, заглянув на мгновение внутрь ячейки, быстро определял скрытую цифру и сразу же двигался к соответствующей следующей ячейке, однако Система не засчитывала это распознавание, если не выполнялось одно из двух условий: скрытая цифра неподвижно зафиксирована на 1 сек.; в поле зрения отсутствуют лишние символы. Поэтому был добавлен программный приём с использованием виртуально-дополненной реальности. В момент, когда в кадре стабильно фиксировались лишь две цифры — номер ячейки и скрытая цифра, — поверх реального изображения автоматически накладывался виртуальный маркер в виде рамки, цвет которой изменялся с красного на зелёный. Смена цвета служила индикатором успешного распознавания и подтверждала возможность перехода к следующей ячейке. Введение этой визуальной обратной связи позволило существенно повысить точность распознавания, приблизив её к 100%.

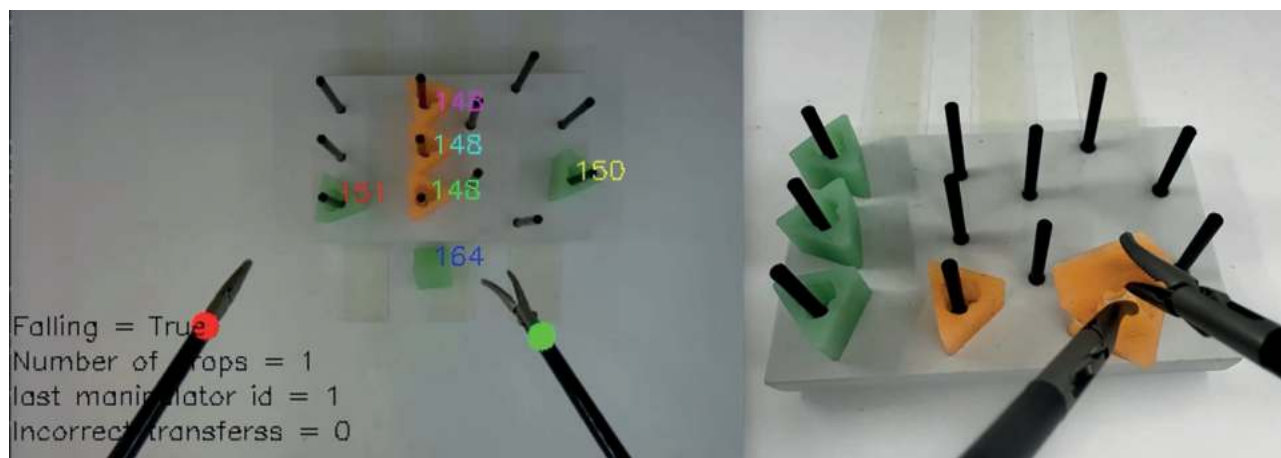


Рис. 5. Определение положения призм — на штырьке, в инструменте или на дне бокса

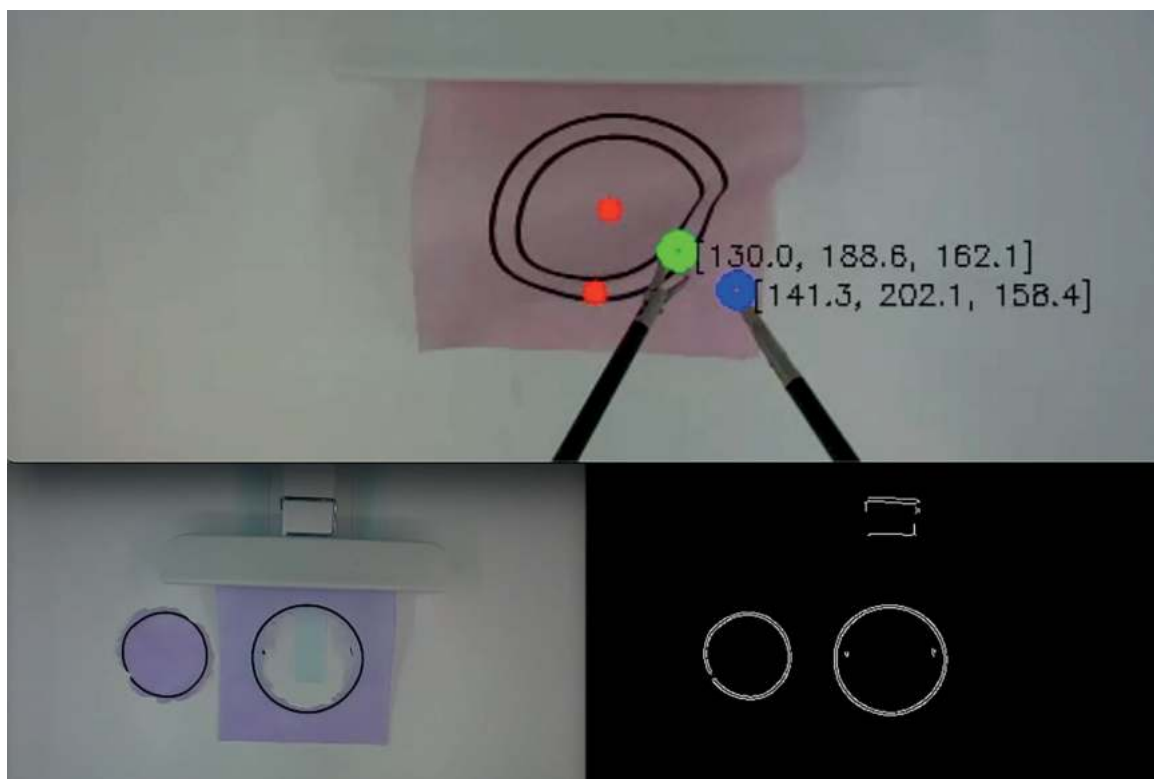


Рис. 6. Обычное и обработанное компьютером изображения с вырезанными кругами

Перемещение объектов. В этом упражнении курсант перекладывает небольшие объекты (силиконовые призмы) с одного штывера на другой с помощью лапароскопических инструментов. Отрабатываемый навык — уверенное бимануальное взаимодействие, ошибки — уроненные объекты. Система БЭСТА-2.0 автоматически подсчитывает число падений призм (рис. 5). Так, на первом этапе ($N = 50$ видео) Система фиксировала до 14 падений объектов вместо двух фактически допущенных, после второго этапа ($N = 170$) погрешность в подсчёте числа падений объектов снизилась до 0–3 (в среднем ~ 1), а после третьего этапа обучения ($N > 300$) для большинства попыток количество падений призм определяется абсолютно верно, и лишь в 3% случаев наблюдается ошибка на одну единицу. Более грубых ошибок с расхождением более чем на одно определение падения на заключительном этапе не отмечалось.

Исечение круга. Применён специальный алгоритм: по цветовой сегментации выделяются тёмные отверстия на фоне (вырезанный круг и отверстие в оставшемся материале), затем с помощью кластеризации DBSCAN и аппроксимации окружностью определяется доля вырезанной окружности (рис. 6). Вычисляется процент завершенности вырезания: 100% соответствует идеально вырезанному кругу. На этапе тестирования было показано, что погрешность измерения этой метрики не превышает $\pm 3\%$ по длине окружности, что сравнимо с точностью ручного измерения. Таким образом, система способна объективно оценить, насколько полно круг вырезан (например, 95% — почти завершено, остался небольшой неотсечённый участок).

Клипирование и пересечение. По ходу упражнения требуется наложить две клипсы на имитацию сосуда (две резинки, растянутые по 12 штыверкам учебной платформы) и затем последовательно пересечь их эндоножницами между клипсами. Правила требуют, чтобы клипсы были наложены на определённом расстоянии, а разрез выполнен ровно посередине между клипсами, двух резинок поочередно. Система технического зрения автоматически определяет моменты наложения каждой клипсы и факт пересечения резинки. Это достигается за счёт комбинированного анализа геометрией формы резинок. После наложения клипсы алгоритм делает «снимок» текущей формы фигур, сформированных двумя резинками, классифицирует их состояние (наложена одна или две клипсы) и сравнивает с эталоном. После пересечения выполняется финальная классификация состояния резинки (пересечена / не пересечена, с клипсами на концах). По результатам валидации система безошибочно определяла правильность выполнения клипирования во всех тестовых видео. В случаях имитации ошибок (например, наложена только одна клипса или клипсы установлены слишком далеко друг от друга) эти отклонения также успешно распознавались, и алгоритм снижал итоговый балл.

После третьего этапа сбора и машинного переобучения ($N > 300$ видео) была достигнута практически полная согласованность автоматизированной оценки Системы с экспертными заключениями. Было подтверждено, что фиксация всех необходимых действий и наличия/отсутствия разрывов производится корректно в подавляющем большинстве случаев.

Также была решена проблема преждевременного автоматического завершения упражнения. По умолчанию считалось, что 10-секундное отсутствие движения в боксе свидетельствует о том, что выполнение задания закончено. Поскольку при данном упражнении инструменты постоянно извлекаются из бокса для перезарядки клипсы, то в ходе тестирования первоначальная версия алгоритма некорректно интерпретировала выход инструмента из кадра как окончание упражнения. Изначальная величина в 10 сек. для данного упражнения была увеличена до 20 сек., что устранило проблему преждевременного финиша при исчезновении инструментов из кадра или фиксации состояния резинки с клипсами на любом этапе.

Наложение эндопетли. Упражнение заключается в затягивании узла эндоскопической петли типа Рёдера на заданной отметке. Ключевой параметр успеха — точность размещения петли на целевой метке. Наша система анализирует финальный кадр: определяет координаты узла петли и метки, измеряет расстояние между ними. По первоначальным испытаниям ошибка в определении этой дистанции достигала 3 см, особенно если нить и метка были схожего цвета (чёрная нить на чёрной полоске). Использование белой нити значительно облегчает задачу, однако современный шовный материал для контрастной визуализации в ране, как правило, имеет тёмный цвет (рис. 7). По мере обучения определение отклонения положения узла петли от целевой метки сократилось с нескольких сантиметров после первого этапа до 0,7 мм (порядка одного пикселя изображения), что фактически равно пределу разрешающей способности видеокамеры.

При использовании петли из чёрной нити на чёрной метке алгоритм неуверенно определял что есть что, поэтому было принято решение рекомендовать ис-

пользовать светлый шовный материал. Тестирование показало, что отклонение петли более 2–3 мм практически не встречается, что соответствует высоким стандартам выполнения. Таким образом, система способна объективно зафиксировать, затянута ли петля поверх метки или имеется расстояние от неё, объективно оценив расхождение вплоть до субмиллиметровых отклонений.

Результаты внедрения системы автоматизированной оценки на основе ИИ в лапароскопическую Систему БЭСТА-2.0 демонстрируют принципиальную реализуемость объективной автоматической оценки эндохирургических навыков (рис. 8). Достигнутая высокая степень совпадения с оценками экспертов (порядка 95–100% на итоговой выборке) свидетельствует о том, что грамотно обученные модели компьютерного зрения могут воспроизводить экспертную оценку действий хирурга. Особенно показательно, что по ключевым параметрам (число ошибок, точность позиционирования узла, полнота выполнения задания) система после обучения на нескольких сотнях примеров выходит на погрешность, сопоставимую с вариативностью между разными экспертами. Таким образом, можно говорить о подтверждении гипотезы о применимости ИИ для автоматизации оценки базовых навыков эндохирурга.

В процессе разработки были выявлены важные особенности. Во-первых, качество алгоритмов значительно зависит от объёма и разнообразия обучающих данных. Начальные версии, обученные на десятках видео, были недостаточно точны, однако по мере накопления данных (до > 300 видео) показатели резко улучшились. Это согласуется с общими принципами машинного обучения — нейросети требуют большого числа примеров для обучения разнообразным ситуациям.

Рис. 7. Наложение эндоскопической петли и затягивание узла точно по метке на отростке



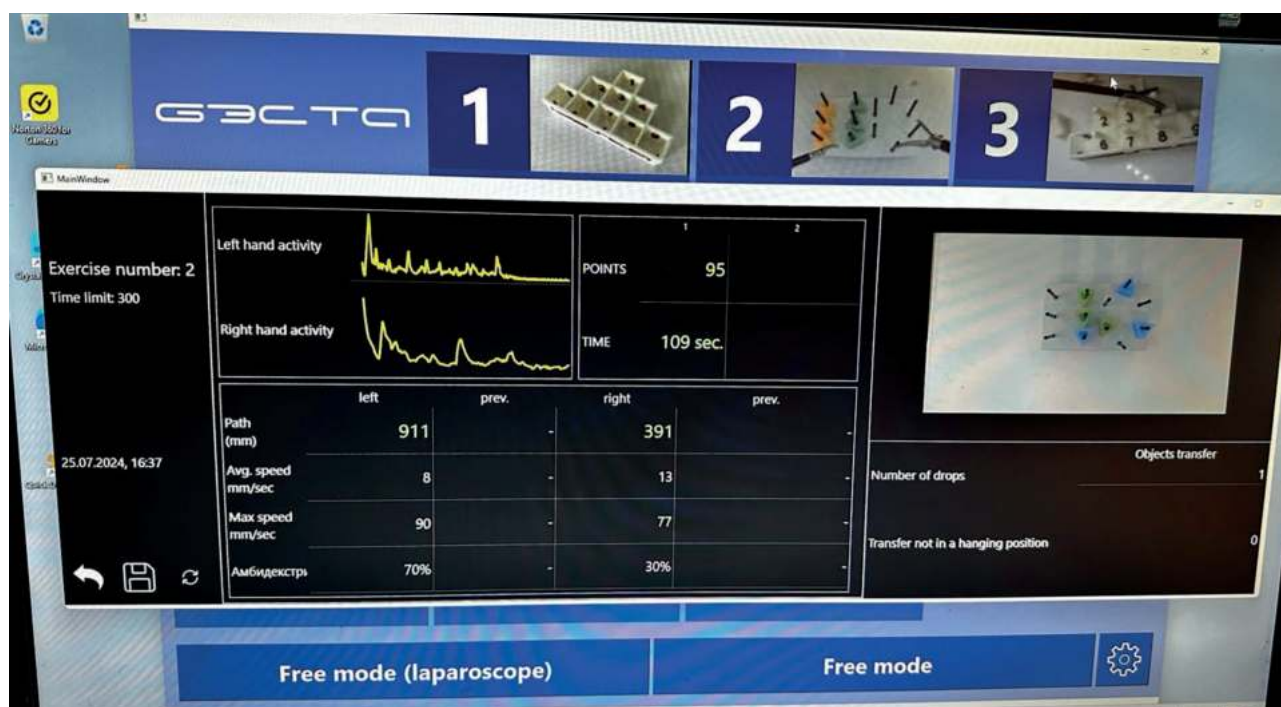


Рис. 8. Пример экрана оценки с автоматически определёнными параметрами и кривыми активности обеих рук

Во-вторых, оказалось необходимым тщательно подбирать и донастраивать метрики оптимизации. Например, вычисление средних абсолютных показателей не всегда приводило к наилучшему практическому результату — важнее было минимизировать число случаев с ошибками оценки, отфильтровать нехарактерные пиковые значения скоростей, артефакты перемещения инструментов. Мы изменили функцию качества таким образом, чтобы она отражала долю полностью правильных оценок, что позволило алгоритму избежать систематических мелких ошибок ценой редких выбросов, что в контексте обучения предпочтительнее. В-третьих, пришлось учитывать или изменять физические и оптические факторы: освещённость, угол наклона камеры, цвет шовного материала, конструкцию оптической системы.

Также выяснилось, что для надёжной работы Системы необходимо обязательно придерживаться стандартизованных правил выполнения упражнений. По итогам испытаний были внесены изменения в регламент выполнения: уточнены описания некоторых упражнений (чтобы исключить неоднозначные ситуации для распознавания), установлен более конкретный подбор расходных материалов, обеспечена стабильная освещённость и размечены границы активной зоны в кадре, чтобы курсанты правильно располагали пособия в боксе и не выходили при манипуляциях за поле зрения 3Dкамеры. Также выявлено, что высота бокса влияет на обзор — было решено увеличить высоту для улучшения фокусировки на инструментах. Эти изменения позволили практически устранить случаи, когда Система ошибалась из-за внешних факторов.

Сравнивая наш подход с известными разработками, отметим, что на мировом рынке начинают появлять-

ся коммерческие тренажёры с элементами AI-оценки навыков [5]. Однако открытых данных об их точности пока мало. В опубликованном в 2022 году исследовании сообщается о создании алгоритма оценки лапароскопических упражнений с согласованностью с экспертами около 85% [4]. Наша система достигла ~90-95% по основным метрикам, а по некоторым из них — все 100% что может объясняться наличием большого количества обучающих видео и стандартизацией выполнения упражнений.

Особенностью нашего подхода является высокая трудоёмкость. Если бы заказчики (преподаватели-хирурги) ограничивались бы оценкой только определения времени, выполнения и траектории (с производными от них — скорость, амбидекстрия), то анализ всех заданий вёлся бы с помощью одного единственного алгоритма. В нашем же случае определялись разные события и параметры по каждому упражнению, для чего потребовались ручная разметка кадров, вслед за которой разметка видео для оценки упражнения, затем обучение сети, разработка отдельного алгоритма под каждое упражнение, тюнинг гиперпараметров — всё это оказалось очень трудоёмким процессом.

Ограничением текущего исследования является то, что оценивались преимущественно базовые навыки на отработанных сценариях упражнений в симулированных условиях бокса-тренажера. Реальные операции гораздо сложнее и вариативнее, и перенос алгоритмов в реальную хирургию потребует решения дополнительных задач (например, распознавания анатомических структур, учёта непредвиденных ситуаций). И тем не менее отработка базовых навыков — необходимый фундамент, и представленный подход может быть расширен для более сложных сценариев.

В будущем, по мере усложнения задач (например, оценка продвинутых оперативных приёмов), возможно снижение точности, и потребуются дополнительно совершенствовать модели и расширять датасеты. Тем не менее полученные результаты обнадеживают: даже базовые модели на уровне детекторов объектов и простых классификаторов способны эффективно решать задачу оценки навыков, если они правильно интегрированы в систему и обучены на репрезентативных данных.

Заключение

Разработана и внедрена система технического зрения БЭСТА-2.0 для автоматизированной оценки базовых навыков эндохирурга. Она включает аппаратно-программный комплекс с лапароскопическим тренажёром, оснащённым стереокамерой, и набор алгоритмов ИИ для распознавания действий и ошибок. Система прошла многоэтапную валидацию: на более чем 300 видеозаписях упражнений показано, что автоматическая оценка совпадает с оценкой экспертов примерно в 95% случаев по ключевым показателям точности (падение объектов, точность наложение эндопелти, последовательность наложения клипс, точность иссечения круга). Продемонстрирована возможность одновременно с распознаванием событий определять и количественные метрики (время, траектории, скорость, симметрия работы руками) для комплексной характеристики уровня выполнения упражнения.

Практическое значение работы состоит в том, что подобная система может снять часть нагрузки с инструкторов, обеспечивая объективный и быстрый анализ действий обучаемого, выявление типичных ошибок и отслеживание прогресса. Автоматизированная платформа стандартизирует оценку и делает её независимой от человеческого фактора, помогая в условиях, когда у преподавателя нет возможности часами наблюдать за тренингом каждого студента и детально разбирать выполнение каждого подхода.

Таким образом, платформа БЭСТА-2.0 закладывает основу для широкого применения ИИ в обучении эндохирургов, помогая оптимизировать хирургическое обучение, повысить эффективность тренинга и безопасность последующих оперативных вмешательств молодых специалистов.

Благодарности

Авторы статьи выражают огромную благодарность нашим коллегам и студентам за их активную исследовательскую и практическую работу. В разработке, создании и апробации Системы приняли участие и внесли важный вклад: А. В. Аракелов, Д. М. Грибков, Д. С. Кучерявый, Д. В. Малоросиянцев, В. Г. Морозов, Л. А. Отдельнов, А. С. Смирнов, А. Ю. Чиндясов, Л. Б. Шубина.

Вклад авторов

Горшков М. Д.: разработка учебно-методической концепции БЭСТА, технического задания.

Горшков М. Д., Ким Е.В., Бесчастнов И.В.: концепция и дизайн исследования; сбор, анализ и интерпретация данных, тестирование Системы, текст статьи.

Бесчастнов И.В., Тюрин Р.Р., Павлов С.А.: программирование и отладка технического зрения.

Все авторы принимали участие в обсуждении и редактировании текста рукописи.

Все авторы утвердили окончательную версию статьи.

Заявление об использовании ИИ

В работе над данным текстом генеративный ИИ был использован в следующих целях: 1. Поиск и реферирование источников информации в интернет; 2. Проверка текста.

ЛИТЕРАТУРА

1. Горшков М. Д., Совцов С. А., Матвеев Н. Л. Допуск ординаторов в эндохирургическую операционную. Какие базовые лапароскопические навыки можно освоить на доклиническом этапе — вне операционной? // Эндокопическая хирургия. 2016. № 1. С. 38–45. doi: 10.17116/endoskop201622138-45
2. Горшков М. Д., Свистунов А. А., Совцов С. А., Матвеев Н. Л., Колыш А. Л. БЭСТА. Базовый эндохирургический симуляционный тренинг и аттестация. 2-е изд., доп. и перераб. М.: РОСМЕД, 2018. 80 с. : ил.
3. Aggarwal R., Grantcharov T., Moorthy K., Hance J., Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. Am J Surg. 2006, Jan. No. 191 (1). P. 128–133. DOI: 10.1016/j.amjsurg.2005.10.014. PMID: 16399123
4. Belmar F., Gaete M. I., Escalona G., Carnier M., Durán V., Villagrán I., Asbun D., Cortés M., Neyem A., Crovari F., Alseidi A., Varas J. Artificial intelligence in laparoscopic simulation: a promising future for large-scale automated evaluations. Surg Endosc. 2023, Jun. No. 37 (6). P. 4942–4946.
5. Ethicon (J&J). AI-powered Laparoscopic Skills Training Platform Debut — Press Release [Электронный ресурс]. 2023, Nov. 6. URL: <https://www.jnj.com/media-center/press-releases/ethicon-debuts-an-ai-powered-laparoscopic-skills-training-platform-at-the-american-association-of-gynecological-laparoscopists-global-congress> (дата обращения: 15.05.2025).
6. Schlosser K., Alkhawaga M., Maschuw K. et al. Training of laparoscopic skills with virtual reality simulator: a critical reappraisal of the learning curve // Eur Surg. 2007. No. 39. P. 180–184. DOI: 10.1007/s10353-006-0292-2
7. Seymour N. E., Gallagher A. G., Roman S. A., O'Brien M. K., Bansal V. K., Andersen D. K., Satava R. M. Virtual reality training improves operating room performance: results of a randomized, double-blinded study // Ann Surg. 2002, Oct. No. 236 (4). P. 458–463; discussion 463–464. DOI: 10.1097/00000658-200210000-00008