

УДК 004.93

DOI: 10.18413/2518-1092-2025-10-3-0-5

Щетинин Е.Ю.¹
Тютюнник А.А.²

О СЕГМЕНТАЦИИ ПОЛИПОВ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ SEGMENT ANYTHING MODEL

¹) Севастопольский государственный университет,
ул. Университетская, 33, г. Севастополь, 299053, Россия
²) Российский университет дружбы народов им. Патриса Лумумбы,
ул. Миклухо-Маклая, 6, Москва, 117198, Россия

e-mail: riviера-molto@mail.ru, tyutyunnik_aa@pfur.ru

Аннотация

Полипы толстой кишки являются ключевыми предикторами развития колоректального рака, и их своевременное выявление играет решающую роль в профилактике онкологических осложнений. В работе предложена усовершенствованная модель автоматической сегментации полипов Polyps-SAM2, основанная на фундаментальной архитектуре Segment Anything Model 2 (SAM2). Модель адаптирована для медицинской визуализации путём тонкой настройки с заморозкой параметров кодировщика изображений и внедрением обучаемых слоёв для обработки текстовых инструкций. Эксперименты проведены на двух общепринятых наборах данных – Kvasir-Seg и CVC-ClinicDB. Polyps-SAM2 продемонстрировала высокую точность: значения метрик Dice и IoU составили 0.94 и 0.91 на Kvasir-Seg, а также 0.938 и 0.901 на CVC-ClinicDB, что превосходит или сопоставимо с современными методами сегментации. Несмотря на ограничения при обработке изображений с множественными полипами и зависимость от подсказок (например, ограничивающих рамок), предложенная модель обладает высокой обобщающей способностью и потенциалом для интеграции в клинические системы поддержки принятия решений при проведении колоноскопии.

Ключевые слова: полипы толстой кишки; сегментация; глубокое обучение; фундаментальные модели компьютерного зрения; трансформер

Для цитирования: Щетинин Е.Ю., Тютюнник А.А. О сегментации полипов с использованием модели Segment Anything Model // Научный результат. Информационные технологии. – Т.10, №3, 2025. – С. 55-63. DOI: 10.18413/2518-1092-2025-10-3-0-5

Shchetinin E.Yu.¹
Tiutyunnik A.A.²

ON SEGMENTATION OF POLYPS USING SEGMENT ANYTHING MODEL

¹) Sevastopol State University,
33 Universitetskaya St., Sevastopol, 299053, Russia

²) RUDN University,
Miklukho-Maklaya St., 6, Moscow, 117198, Russia

e-mail: riviера-molto@mail.ru, tyutyunnik_aa@pfur.ru

Abstract

Colorectal polyps are critical precursors to colorectal cancer, and their early detection is vital for effective prevention. This study introduces Polyps-SAM2 – an enhanced polyp segmentation model built upon the Segment Anything Model 2 (SAM2) foundation. Tailored for medical imaging, Polyps-SAM2 incorporates fine-tuning with a frozen image encoder and integrates trainable layers for processing textual prompts. Evaluated on two benchmark datasets – Kvasir-Seg and CVC-ClinicDB – the model achieves Dice and IoU scores of 0.94/0.91 and 0.938/0.901, respectively, outperforming or matching state-of-the-art segmentation approaches. While limitations remain – particularly in handling images with multiple distinct polyps and reliance on user-provided prompts such as bounding boxes – the model demonstrates strong generalization capabilities and significant potential for clinical deployment in computer-aided colonoscopy systems, thereby improving diagnostic accuracy and workflow efficiency for physicians.

Keywords: colon polyps; segmentation; deep learning; fundamental computer vision models; transformer

For citation: Shchetinin E.Yu., Tiutiunnik A.A. On Segmentation of Polyps Using Segment Anything Model // Research result. Information technologies. – Т.10, №3, 2025. – P. 55-63. DOI: 10.18413/2518-1092-2025-10-3-0-5

ВВЕДЕНИЕ

Колоректальный рак (КРР) – распространенная злокачественная опухоль желудочно-кишечного тракта с высокой заболеваемостью и низкой выживаемостью при поздней диагностике. КРР является второй ведущей причиной смертности от рака в мире: в 2022 году будет зарегистрировано более 106 180 случаев заболевания и 52 580 смертей [1]. Наличие полипов толстой кишки является важным фактором в возникновении КРР. Раннее выявление и удаление полипов может значительно снизить смертность и улучшить результаты лечения пациентов [2]. Колоноскопия считается золотым стандартом диагностики КРР. Она позволяет выявить и устранить полипы толстой кишки, что может предотвратить дальнейшее повреждение окружающих тканей и снизить вероятность развития КРР.

В настоящее время скрининг полипов толстой кишки основывается на осмотре эндоскопических изображений вручную. Однако этот метод является достаточно трудоемким, а также подвержен ошибкам врача. Использование компьютерной диагностики для колоноскопии может эффективно повысить эффективность аннотации и сократить время постановки диагноза КРР [3]. В последние годы методы, основанные на глубоком обучении, достигли значительного прогресса в сегментации полипов. Такие модели, как U-Net [4] и UNet++ [5], продемонстрировали эффективность сверточных нейронных сетей в извлечении пространственных признаков для точной локализации и сегментации полипов. Появление архитектур нейронных сетей на основе трансформеров, в частности Vision Transformer (ViT), продвинуло исследования этой области еще дальше благодаря использованию механизмов внутреннего внимания для выявления пространственных зависимостей и контекстной информации. Однако, несмотря на определенные достижения на эталонных наборах данных, многие из этих моделей демонстрируют ограниченную способность к адаптации при использовании их на небольшом частном наборе данных, что подчеркивает необходимость создания более адаптируемой и надежной системы сегментации производительность.

Фундаментальные модели глубокого обучения в последнее время привлекают внимание в задачах компьютерного зрения благодаря своим возможностям к обобщению в различных задачах компьютерного зрения, включая классификацию изображений, обнаружение объектов и сегментацию. Так, передовая фундаментальная модель Segment Anything Model (SAM) продемонстрировала впечатляющие возможности сегментации благодаря механизму встроенных подсказок и инструкций (prompts) [6,7]. Современные модели все чаще используют возможности включения различных модальностей, например, текстовых подсказок и инструкций к изображениям. Основываясь на этом, современные модели, такие как SAM2 [8], еще более усовершенствовали процесс сегментации. Модель SAM2 использует функцию Memory Attention и выигрывает от обучения на больших наборах данных, улучшая сегментацию видео и изображений. Однако производительность SAM2 в таких специализированных областях, как медицинская визуализация, по-прежнему остается неоптимальной [9].

В настоящей работе использованы фундаментальные модели компьютерного зрения, предварительно обученные на крупнейших базах знаний, а также точной настройки блока кодировщика модели SAM2 на нескольких наборах изображений полипов Kvasir-Seg и CVC-ClinicDB. Основной вклад данного исследования в разработку систем автоматизированного процесса сегментации полипов состоит в следующем:

1. Предложена мультимодальная модель Polyps-SAM2 для решения задачи сегментации полипов по изображениям из наборов Kvasir-SEG, CVC-ClinicDB. Для этого была выполнена тонкая настройка модели SAM2, на этапе обучения заморозив параметры кодировщика в архитектуре SAM2 и встроив в нее несколько обучаемых слоев текстовых инструкций.

2. Проведенные в исследовании численные эксперименты показали, что предложенная модель превосходит существующие современные методы сегментации по ряду показателей качества сегментации.

МЕТОДЫ ИССЛЕДОВАНИЯ

Архитектура модели SAM

SAM – фундаментальная модель компьютерного зрения, которая сегментирует заданные пользователем объекты интереса с помощью его подсказок, таких как точка, маска, ограничивающая рамка или текст [7]. Модель SAM состоит из трех компонентов: кодировщика изображения, кодировщика подсказки (prompts) и декодера маски. В модели SAM в качестве кодировщика изображений использован трансформер ViT, модифицированный для обработки входных изображений с разрешением до 1024x1024 пикселей. Трансформер ViT состоит из ряда слоев внутреннего внимания для выявления пространственных зависимостей между различными участками входного изображения. Затем следует набор слоев прямого прохождения, которые преобразуют выход слоев внутреннего внимания в набор карт признаков, используемых в дальнейшем декодере для создания масок сегментации. Кодировщики инструкций разработаны для подсказок пользователем, таких как точки, поля, тексты и маски. Ограничительная рамка кодируется точечным кодированием ее левого верхнего угла и правого нижнего угла. Текст в свободной форме кодируется предварительно обученным текстовым кодировщиком CLIP. Маска подсказки имеет то же пространственное разрешение, что и входное изображение, которое кодируется сверточными картами признаков. Наконец, декодер маски использует архитектуру, состоящую из двух слоев трансформера с динамической головкой предсказания маски и головкой регрессии оценки пересечения над объединением (IoU). Головка предсказания маски может генерировать уменьшенные маски, которые соответствуют целому объекту, части объекта. На Рис.1 представлена архитектура модели сегментации SAM.

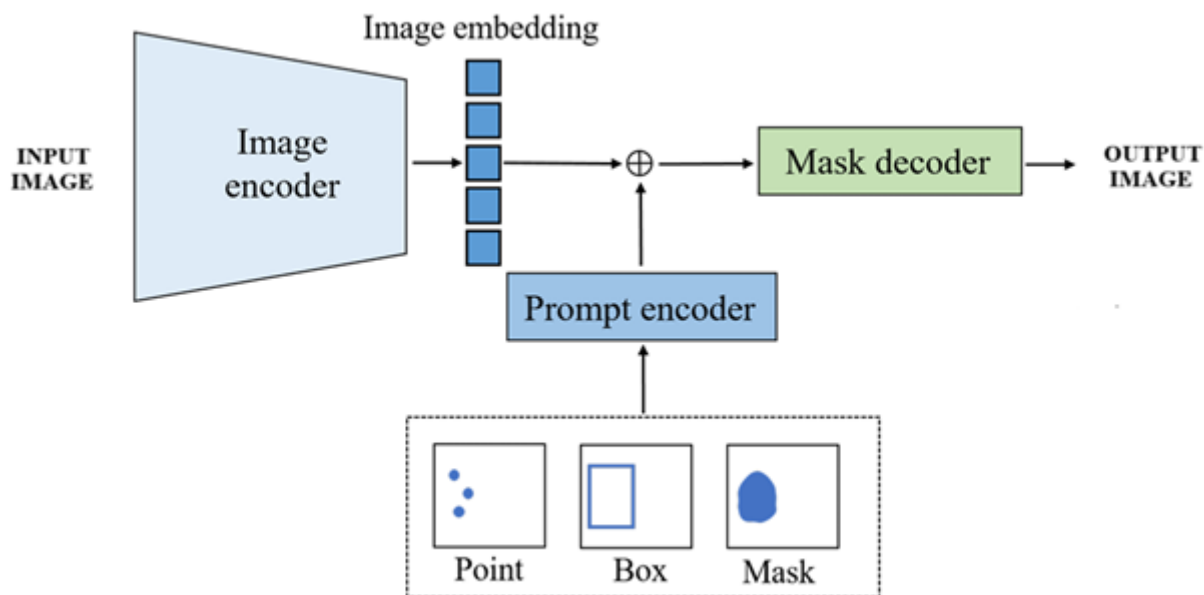


Рис. 1. Архитектура модели Segment Anything Model, которая состоит из кодировщика ViT для извлечения эмбедингов изображений, кодер подсказок для интеграции взаимодействия с пользователем через различные режимы подсказок и декодер масок для предсказания масок сегментации путем объединения эмбедингов изображений и эмбедингов подсказок [6]

Fig. 1. Architecture of the Segment Anything Model, which consists of a ViT encoder for extracting image embeddings, a sub-tip encoder for integrating user interaction through different hint modes, and a mask decoder for predicting segmentation masks by combining image embeddings and hint embeddings [6]

Архитектура модели Segment Anything Model 2 (SAM2)

SAM2 – фундаментальная модель компьютерного зрения, разработанная для решения задачи сегментации на изображениях и видео [8]. В этой версии модели использована архитектура трансформера, включающая в себя дополнительные компоненты, такие как кодировщик памяти, банк памяти и модуль внимания к памяти, который хранит и использует информацию об объектах. Эти усовершенствования позволяют эффективно обрабатывать и использовать информацию из памяти для работы с видео в реальном времени. Будучи развитием и продолжением модели сегментации SAM, SAM2 имеет множество улучшений, значительно повышающих его возможности сегментации:

Возможность сегментации видео: В SAM2 реализована поддержка сегментации видео, что позволяет сегментировать Объекты в видео и выполнять кросс-кадровое отслеживание и редактирование.

Повышенная точность и скорость: при выполнении одних и тех же задач по сегментации изображений SAM2 достигает шестикратного по сравнению с SAM при сохранении высокой точности сегментации.

Мелкозернистая сегментация: SAM2 обеспечивает более точную и тонкую сегментацию, извлекая более глубокую семантическую информацию, что повышает его потенциал для тонкой настройки в задачах сегментации мелких объектов по сравнению с SAM.

ЭКСПЕРИМЕНТЫ

Наборы данных

Для тестирования предложенной модели и тонкой настройки мы провели сравнительные эксперименты на пяти эталонных наборах данных колоноскопии:

Kvasir-SEG [9] – набор данных создан компанией Vestre Viken Health Trust (Норвегия) и состоит из 1 000 изображений полипов и соответствующих им меток, полученных на основе анализа видеопоследовательностей результатов колоноскопии. Изображения были вручную аннотированы врачами, а затем проверены опытными гастроэнтерологами. Разрешение изображений варьируется от (332×487) до (1920×1072) пикселей. Набор данных предназначен для исследований и разработки новых и усовершенствованных методов сегментации, обнаружения, локализации и классификации полипов, предлагая передовые решения для задач, связанных с полипами.

CVC-ClinicDB [10] – набор данных медицинской визуализации, предназначенный для выявления и исследования колоректального рака, в частности, для раннего выявления и диагностики колоректального рака. Он состоит из 612 снимков высокого разрешения изображения колоноскопии из 29 видеопоследовательностей колоноскопии с разрешением 288×384, полученные от пациентов с раком прямой кишки, причем каждое изображение было аннотировано экспертами, чтобы отметить раковые поражения. Этот набор данных широко используется для обучения и проверки моделей колоректального рака, сегментации изображений, обработки медицинских изображений и клинических исследований. Набор данных включает 612 исходных изображений с соответствующими базовыми метками сегментации, каждое из которых представляет области, покрытые полипами. Разрешение изображений составляет (384×288×3). На Рис. 2 показаны примеры изображений полипов с их базовыми истинами для сегментации из набора данных CVC-ClinicDB.

Метрики оценки производительности модели сегментации

Производительность моделей сегментации оценивалась с помощью нескольких метрик качества сегментации Dice, IoU (Intersection over Union). Показатель Dice дает сбалансированное представление о производительности модели, измеряя степень совпадения между предсказанными масками и истинными масками. Показатель IoU оценивает пространственное совпадение между предсказанными изменениями и реальными. Эти метрики определяются следующим образом:

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN}, \quad (1)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (2)$$

where TP, FP, FN, and TN are the true positive, false positive, false negative, and true negative counts, respectively.

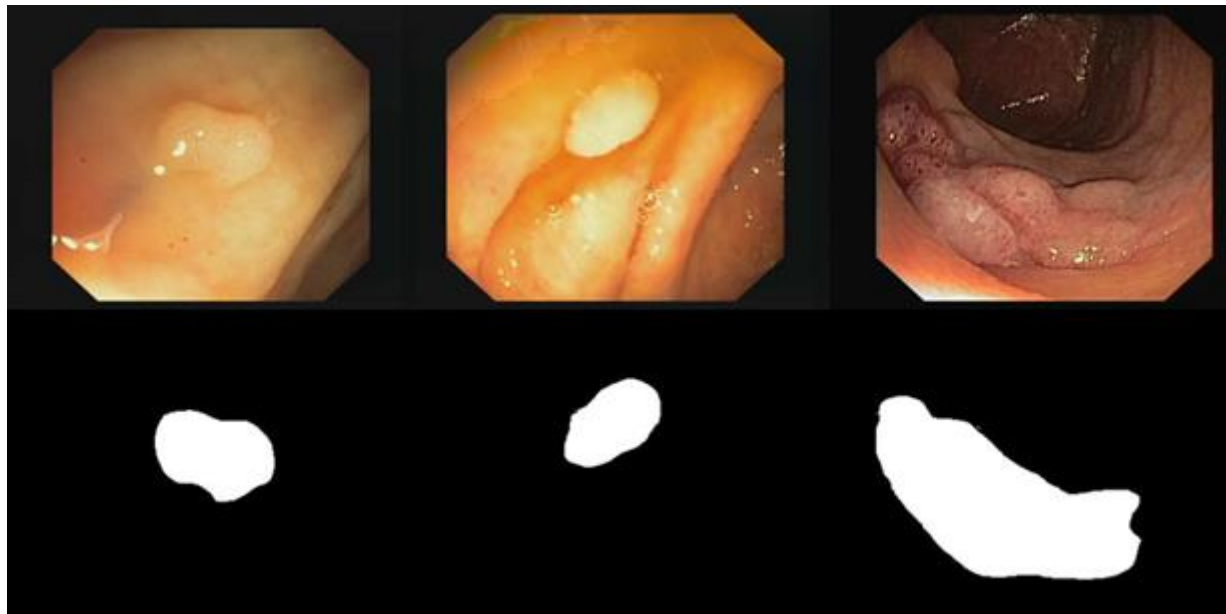


Рис. 2. Образцы изображений полипов (верхний ряд) и истинные маски сегментации, соответствующие областям, содержащим полипы (нижний ряд), полученных из набора данных CVC-ClinicDB

Fig. 2. Sample polyp images (top row) and true segmentation labels corresponding to areas containing polyps (bottom row) obtained from the CVC-ClinicDB dataset

Процесс обучения модели Polyps-SAM2

В модели Polyps-SAM2 в качестве кодировщика изображений использован трансформер ViT-Base. Все эмбединги изображений вычислялись в автономном режиме путем предоставления нормализованных изображений кодировщику, который затем изменял их размеры до $(3 \times 1024 \times 1024)$. В процессе обучения было применено несколько методов увеличения количества изображений. Во-первых, использовано гауссово размытие к изображениям. Затем случайным образом изменялись яркость, контрастность и оттенок изображения. Изображения также подвергались горизонтальному и вертикальному перевороту. Кроме того, были применены аффинные преобразования включая масштабирование и сдвиг. Гауссово размытие и цветовое дрожание применялись только к изображениям, в то время как переворачивание и аффинные преобразования применялись как к изображениям, так и к соответствующим им маскам сегментации.

Модель Polyps-SAM2 построена с использованием фреймворка PyTorch и обучена на одном графическом процессоре NVIDIA RTX 3070 с 12 ГБ VRAM. Во время обучения использован метод оптимизации AdamW с начальной скоростью обучения $1E-4$, размером партии $batch_size=4$ и общим количеством 200 эпох. Функция потерь сочетала в себе потери по показателю Dice и бинарной кросс-энтропии. Планировщик скорости обучения уменьшал скорость обучения в два раза, если оценка по показателю Dice не улучшалась, при этом минимальная скорость обучения составляла $1E-5$. Производительность модели оценивалась с помощью таких показателей, как Dice, IoU, precision, recall и accuracy, и модель сохранялась, когда показатель Dice улучшался.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В работе проведен сравнительный анализ результатов применения модели Polyps-SAM2 и некоторых известных моделей сегментации полипов на наборах Kvasir-SEG и CVC-ClinicDB. Для этого было выбрано несколько моделей таких как U-Net [4], UNet++ [5], PraNet [11], EU-Net [12], HSNet [13], MSRF-Net [14], Polyp PVT [15], Polyp SAM [16], MSEG [17]. В таблице представлены результаты сравнительного анализа модели Polyps-SAM2 с остальными указанными выше моделями на наборах изображений для сегментации полипов Kvasir-SEG, CVC-ClinicDB. По итогам сравнительного анализа модель Polyps-SAM2 продемонстрировала свою высокую эффективность в задаче сегментации полипов.

Таблица

Результаты сравнительного анализа моделей сегментации на наборах Kvasir-SEG, CVC-ClinicDB
Table

Results of comparative analysis of segmentation models on Kvasir-SEG, CVC-ClinicDB sets

Model	Kvasir-SEG		CVC-ClinicDB	
	Dice	IoU	Dice	IoU
UNet [3]	0.78	0.731	0.77	0.73
Unet++[4]	0.81	0.78	0.80	0.76
PraNet [11]	0.84	0.81	0.88	0.79
EU-Net[12]	0.89	0.84	0.88	0.804
HSNet [13]	0.915	0.88	0.904	0.848
MSRF-Net [14]	0.87	0.79	0.86	0.79
MSEG [17]	0.919	0.89	0.907	0.85
Polyp PVT [15]	0.927	0.901	0.91	0.87
Polyp SAM [16]	0.92	0.904	0.912	0.88
Polyps-SAM2	0.94	0.91	0.938	0.901

Из Таблицы следует, что по обоим показателям Dice, IoU она превосходит остальные модели сегментации полипов, использованные в данном исследовании. Модели Polyp SAM и Polyp PVT также достигли выдающихся результатов в решении сегментации полипов при относительно невысоких затратах вычислительных ресурсов, тогда как для тонкой настройки модели Polyps-SAM2 потребовалось значительно большие временные и аппаратные ресурсы.

На рис. 3 приведены результаты применения модели Polyps-SAM2 к сегментации полипов по изображениям из набора CVC-ClinicDB dataset. На изображении слева представлен снимок полипа и его истинная маска. Справа приведена маска, предсказанная моделью Polyps-SAM2.



а) Изображение полипа и его истинная маска

б) Предсказанная маска
моделью Polyps-SAM2

Рис. 3. Пример сегментации полипа на изображении из набора Kvasir-SEG моделью Polyps-SAM2
Fig. 3. Example of polyp segmentation in an image from the Kvasir-SEG set by the Polyps-SAM2 model

Также, несмотря на полученные высокие результаты, предложенная модель Polyps-SAM2 имеет некоторые ограничения. Как показано на рисунке 4, она не справляется с множественными отдельными полипами на изображениях колоноскопии. В планах данного исследования попытаться решить эту проблему можно путем улучшения настройки декодеров масок для вывода нескольких масок сегментации. Кроме того, поскольку алгоритм SAM2 основан на изучении визуальных паттернов по подсказкам пользователя, что является основным узким местом для разработки полностью автоматизированных систем сегментации полипов с помощью SAM2.

Тем не менее несмотря на то, что для оптимальной работы модели Polyps-SAM2 в качестве подсказок требуются ограничивающие рамки полипов, можно утверждать, ее применение позволит значительно повысить качество работы врачей и повышает точность обнаружения полипов. Кроме того, PolypsSAM2 является моделью сегментации изображений, в то время как клинические приложения для колоноскопии должны быть моделями сегментации видео.



Рис. 4. Пример обработки изображения моделью Polyps-SAM2, содержащего множественные полипы. Слева исходный снимок множественных полипов. Средний снимок описывает его маску.

Справа снимок является маской полученной моделью Polyps-SAM2

Fig. 4. Example of image processing by Polyps-SAM2 model containing multiple polyps. On the left is the original image of multiple polyps. The middle image describes its mask. The right image is the mask obtained by Polyps-SAM2 model

ЗАКЛЮЧЕНИЕ

В данной работе представлена адаптированная модель Polyps-SAM2 – специализированная версия фундаментальной архитектуры Segment Anything Model 2 (SAM2), нацеленная на задачу сегментации колоректальных полипов по эндоскопическим изображениям. Проведённые эксперименты на двух общепринятых наборах данных – Kvasir-SEG и CVC-ClinicDB – продемонстрировали высокую точность модели: значения метрик Dice и IoU достигли 0.94 / 0.91 и 0.938 / 0.901 соответственно, что сопоставимо или превосходит современные методы, включая Polyp-SAM и Polyp-PVT. Это свидетельствует о высокой обобщающей способности Polyps-SAM2 даже при ограниченном объёме обучающих данных, что особенно ценно в условиях дефицита размеченных медицинских изображений.

Ключевым достоинством предложенного подхода является гибкость архитектуры SAM2, позволяющая интегрировать мультимодальные подсказки (в том числе текстовые инструкции) и эффективно использовать предобученные представления изображений. Кроме того, заморозка параметров кодировщика изображений позволила снизить вычислительные затраты при сохранении высокой семантической выразительности признаков.

Вместе с тем, модель имеет ряд следующих ограничений. Во-первых, её производительность зависит от качества внешних подсказок (например, ограничивающих рамок), что препятствует созданию полностью автоматизированной системы. Во-вторых, как показано на рис. 4, Polyps-SAM2 не справляется с сегментацией множественных отдельных полипов в рамках одного изображения, поскольку декодер маски по умолчанию генерирует только одну маску на подсказку. В-третьих, несмотря на то что SAM2 изначально разрабатывалась как видео-ориентированная модель, в настоящей работе были использованы только статические кадры, что отчасти объясняется необходимостью значительных вычислительных ресурсов.

Таким образом, исходя из сформулированных ограничений, перспективы дальнейшего развития модели Polyps-SAM2 видятся в нескольких направлениях. Интеграция автоматического детектора подсказок (например, на базе YOLOv8 или EfficientDet) позволит устранить зависимость от ручного ввода bounding boxes и приблизить систему к полностью автономному режиму работы. Адаптация к видеоданным с использованием банка памяти SAM2 откроет возможность отслеживания полипов в реальном времени во время колоноскопии, что особенно важно для клинического применения. Модификация декодера маски для поддержки множественных выходов (multi-mask prediction) решит проблему сегментации множественных полипов.

Таким образом, Polyps-SAM2 представляет собой перспективную основу для создания интеллектуальных систем поддержки принятия решений в гастроэнтерологии, однако её переход от исследовательского прототипа к клиническому инструменту требует дальнейшего инженерного и методологического развития.

БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке Севастопольского государственного университета, проект 42-01-09/319/2025-1.

Список литературы

References

1. Siegel R.L., Miller K.D., Wagle N.S., Jemal A. Cancer statistics, 2023. CA: A Cancer Journal for Clinicians. 2023. 73(1), p.17-48, doi: 10.3322/caac.21763
2. Antonelli G., Badalamenti M., Hassan C., Repici A. Impact of artificial intelligence on colorectal polyp detection. Best practice & research. Clinical gastroenterology. 2021. 52-53, 101713, doi: 10.1016/j.bpg.2020.101713
3. Shchetinin E.Yu., On the effectiveness of using visual transformers in detecting abnormalities of histopathological images. Mathematical models and computer simulations. 2025. 17(1), p.46-54, doi: 10.1134/S2070048224700716
4. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science. 2015. V. 9351, doi: 10.1007/978-3-319-24574-4_28
5. Zhou Z., Rahman Siddiquee M.M., Tajbakhsh N., Liang J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA ML-CDS 2018. Lecture Notes in Computer Science. 2018. V. 11045, doi: 10.1007/978-3-030-00889-5_1
6. Kirillov A., et al., Segment Anything. 2023. URL: <https://arxiv.org/abs/2304.02643>
7. He S., et al., Accuracy of Segment-Anything Model (SAM) in medical image segmentation tasks. 2023. URL: <https://arxiv.org/abs/2304.09324>
8. Dong H., Gu H., Chen Y., Yang J., Mazurowski, M. Segment anything model 2: an application to 2d and 3d medical images. 2024. URL: <https://arxiv.org/abs/2408.00756>
9. Jha D. et al. Kvasir-SEG: A Segmented Polyp Dataset. MultiMedia Modeling. MMM 2020. Lecture Notes in Computer Science. 2020. V. 11962, doi: 10.1007/978-3-030-37734-2_37
10. De La Torre R.M., Gonzalez A.A.C., Romero-Aroca A. CVC-ClinicDB: A public database for ophthalmology research. Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2018). 2018. pp. 430-439, doi: 10.1007/978-3-030-00934-2_49.

11. Fan D-P, et al. PraNet: Parallel Reverse Attention Network for Polyp Segmentation. Proceedings of the 2020 International Conference on Medical Image Computing and Computer-Assisted Intervention. 2020. p. 263–273, doi: 10.1007/978-3-030-59725-2_26
12. Patel K., Bur A.M., Wang G. Enhanced U-Net: A Feature Enhancement Network for Polyp Segmentation. IEEE International Conference on Robotics and Vision (ICRV). 2021. pp. 181-188, doi: 10.1109/CRV52889.2021.00032.
13. Zhang W., et al., HSNNet: A hybrid semantic network for polyp segmentation. Computers in Biology and Medicine. 2022. 150, p. 106173, doi: 10.1016/j.combiomed.2022.106173
14. Srivastava A., Jha D., Chanda, S., et al. MSRF-Net: A Multi-Scale Residual Fusion Network for Biomedical Image Segmentation. IEEE Journal of Biomedical and Health Informatics. 2021. 26(5), 2252–2263, doi: 10.1109/JBHI.2021.3138024
15. Dong B., et al., Polyp-PVT: Polyp segmentation with pyramid vision transformers. 2021. URL: <https://arxiv.org/abs/2108.06932>
16. Yuheng L., Mingzhe H., Xiaofeng Ya. Polyp-SAM: Transfer SAM for Polyp Segmentation. 2023. URL: <https://arxiv.org/abs/2305.00293>
17. Chih-Hung H., Hsin-Yu W., Yu-Lun L. HarDNet-MSEG: A Simple Encoder-Decoder Polyp Segmentation Neural Network that Achieves over 0.9 Mean Dice and 86 FPS. 2021. URL: <https://arxiv.org/abs/2101.07172>

Щетинин Евгений Юрьевич, доктор физико-математических наук, доцент, профессор кафедры информационных технологий и систем, Севастопольский государственный университет, г. Севастополь, Россия

Тютюнник Анастасия Александровна, кандидат физико-математических наук, доцент кафедры математического моделирования и искусственного интеллекта, Российский университет дружбы народов им. Патриса Лумумбы, г. Москва, Россия

Shchetinin Eugene Yurievich, Doctor of Physical and Mathematical Sciences, Associate Professor, Professor of the Department of Information Technology and Computer Systems, Sevastopol State University, Sevastopol, Russia

Tiutiunnik Anastasiia Alexandrovna, Candidate of Physics and Mathematics, Associate Professor of the Department of Mathematical Modeling and Artificial Intelligence, RUDN University, Moscow, Russia