Министерство образования Республики Беларусь

Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.932.2

Харкевич
Антон Павлович

Алгоритмы распознавания контуров объектов

АВТОРЕФЕРАТ
на соискание степени магистра

по специальности 7-06-0612-02 – Информатика и технологии программирования

Научный руководитель
Боброва Н.Л.
к.т.н., доцент

Минск 2026

КРАТКОЕ ВВЕДЕНИЕ

Одним из ключевых направлений развития современных технологий компьютерного зрения является повышение точности и надёжности анализа изображений, что, в свою очередь, способствует расширению сфер применения автоматизированных систем обработки визуальной информации. В этих условиях особую значимость приобретает возможность корректного определения состава объектов на изображении, их положения и границ.

Современные нейросетевые модели для обработки изображений представляют собой комплекс алгоритмов и программных средств, обеспечивающих решение основных задач компьютерного зрения в реальном времени. Наиболее популярными задачами компьютерного зрения являются такие задачи, как классификация изображений, обнаружение объектов на изображении и сегментация изображения.

Особое место среди перечисленных задач занимает сегментация экземпляров, сочетающая в себе возможности решения задач обнаружения объектов на изображении и семантической сегментации изображения. Если ранее для решения данных задач требовалась одновременная работа двух независимых моделей, то современные архитектуры позволяют выполнять их в рамках единой модели. Это стало возможным благодаря внедрению методов точного определения контуров объектов, что делает задачу поиска контуров объектов одним из ключевых аспектов современной обработки изображений.

Область нейросетевых моделей, применимых к задачам компьютерного зрения, бурно развивается. В течение последнего десятилетия архитектура моделей и алгоритмы распознавания контуров объектов внутри моделей усложнялись, становясь всё более совершенными. На сегодняшний день обновлённые версии существующих моделей, а также новые модели выходят каждые несколько месяцев. В ходе развития области было создано множество моделей крупными компаниями. Каждая из моделей обладает своими достоинствами и недостатками. Модели различаются по таким показателям, как точность, время работы, возможность дообучения под конкретную область работы. Функциональные характеристики моделей во многом определяются уровнем развития их архитектуры и их программной реализацией.

Диссертационная работа посвящена отслеживанию развития моделей за последнее десятилетие, выявлению ключевых достоинств и недостатков наиболее эффективных моделей на практическом примере, а также созданию гибкого решения в виде программного обеспечения, позволяющего использовать каждую модель в наиболее подходящем для этого случае.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является разработка программного обеспечения для гибкого использования различных нейросетевых моделей, предназначенных для решения задачи распознавания контуров объектов.

Для достижения поставленной цели необходимо решить следующие задачи:

1 Определить основные модели, подходящие для использования в программном обеспечении.

2 Определить набор данных, используемый для всестороннего тестирования моделей.

3 На основе результатов тестирования выявить сильные и слабые стороны каждого из алгоритмов распознавания контуров объектов.

4 Разработать программное обеспечение для гибкого применения лучших из выбранных моделей.

5 Провести проверку разработанного программного обеспечения.

Объектом исследования являются алгоритмы распознавания контуров объектов.

Предметом исследования являются сильные и слабые стороны алгоритмов распознавания контуров объектов.

Основная *гипотеза*, положенная в основу диссертационной работы, заключается в возможности гибкого использования нейросетевых моделей, основанных на различных алгоритмах распознавания контуров объектов. Гибкое использование моделей позволит подчеркнуть достоинства каждого из алгоритмов, лежащих в основе моделей, а также расширить функциональные возможности систем, предназначенных для распознавания контуров объектов.

Практическая значимость работы заключается в возможности применения разработанного программного обеспечения в задачах компьютерного зрения, включая анализ изображений в промышленности, медицине и системах видеонаблюдения.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

Работа выполнялась в соответствии с научно-техническими заданиями и планами работ кафедры информатики.

Личный вклад соискателя

Результаты, приведённые в диссертации, получены соискателем лично. Вклад научного руководителя Н. Л. Бобровой заключается в формулировке целей и задач исследования.

Апробация результатов диссертации

Основные положения диссертационной работы докладывались и обсуждались на XXXIII Международной научно-практической конференции «Актуальные вопросы современной науки и образования» (Пенза, Россия, 2023); XVIII Международной научно-практической конференции «Актуальные научные исследования» (Пенза, Россия, 2024); II Международной научно-практической конференции «Современное общество, наука и образование: актуальные вопросы, достижения и инновации» (Пенза, Россия, 2025); XVII Международной научно-практической конференции «Актуальные вопросы современных научных исследований» (Пенза, Россия, 2025);

Публикации результатов диссертации

По теме диссертации опубликовано 4 работы в сборниках трудов и материалов международных конференций.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, трёх глав, заключения, списка использованных источников, списка публикаций автора и приложений. В первой главе представлен анализ предметной области, выявлены основные архитектурные решения, предназначенные для задачи распознавания контуров объектов, рассмотрены области применения этих решений и их особенности. Вторая глава посвящена тестированию работы выбранных нейросетевых моделей, предназначенных для решения задачи распознавания контуров объектов, а также выявлению основных достоинств и недостатков этих моделей. В третьей главе описывается разработанное программное обеспечение, предназначенное для гибкого использования выбранных моделей, а также производится оценка работоспособности программного обеспечения с помощью наглядных тестов.

Общий объем работы составляет 99 страниц, из которых основного текста –56 страниц, 42 рисунка на 16 страницах, 9 таблиц на 4 страницах, список использованных источников из 42 наименований на 4 страницах и приложения на 19 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

Во **введении** показана актуальность темы диссертационной работы, определены цель и задачи исследования, указаны основные этапы реализации этой цели.

В **первой главе** работы были рассмотрены различные нейросетевые модели, подходящие для выполнения задачи сегментации экземпляров на изображении.

Была рассмотрена каждая версия нейронных сетей семейства R-CNN, начиная от самой первой R-CNN, подходящей только для задачи обнаружения объектов, и заканчивая Mask R-CNN, способной выполнять задачу сегментации экземпляров.

Была рассмотрена модель YOLO, начиная с первой версии и заканчивая последней двенадцатой версией.

Была рассмотрена модель SAM и особенности, которые позволили ей стать фундаментальной моделью.

Была рассмотрена фундаментальная модель следующего поколения SAM 2 и её отличия от модели прошлого поколения.

Была рассмотрена модель OneFormer и её ключевое отличие от других моделей.

Для каждой модели была детально рассмотрена её архитектура, алгоритмы и методы, позволяющие выполнять поставленную задачу (обнаружение объектов или сегментация экземпляров в зависимости от версии). Для каждой новой версии модели были рассмотрены отличия в методах и алгоритмах, улучшающие работу прошлой версии модели, причины и цели разработки этих отличий. Для последней версии каждой модели были приведены её главные достоинства и недостатки.

В результате анализа литературных источников было выявлено следующее:

1 Каждый алгоритм, лежащий в основе рассмотренных моделей, обладает достоинствами и недостатками, которые необходимо проверить на практике на конкретном наборе данных.

2 Не существует идеального алгоритма работы модели, который обходит остальные по всем показателям.

Во **второй главе** работы было произведено тестирование алгоритмов распознавания контуров объектов, рассмотренных в первой главе. Основные этапы процесса подготовки к тестированию и самого тестирования приведены ниже:

1 Был произведён поиск подходящего для тестирования набора данных. В результате был выбран набор данных *BSDS500*.

2 Была произведена подготовка набора данных к тестированию. Изображения набора данных были разделены на категории *Simple, Medium и Hard* в зависимости от сложности, связанной с правильным обнаружением на них контуров объектов.

3 Был произведён выбор моделей прогнозирования, которые будут участвовать в дальнейшем тестировании. В результате было выбрано четыре модели: YOLO, SAM, SAM 2 и OneFormer.

4 Было произведено преобразование результатов прогнозов моделей к формату, удобному для подсчёта метрик.

5 Был произведён выбор метрик для оценки качества прогноза моделей. В качестве основных метрик были выбраны: попиксельная точность (pixel accuracy), SSIM (Structural Similarity Index Measure), IoU (Intersection over Union),

Recall (полнота), Precision (точность), Индекс Дайса (Dice Index или F1-score), Human Rate. Значение каждой метрики сохранялось в файле для дальнейшего тестирования.

6 Был произведён анализ результатов прогнозирования контуров объектов моделью YOLO. Были выявлены основные достоинства и недостатки модели. Был произведён анализ корреляции значений двух наиболее информативных метрик (полнота и Human Rate). Был произведён анализ изменения метрик при переходе от простых к сложным контурам. Было произведено сравнение значений метрик, полученных для модели YOLO, со значениями метрик, полученных для других моделей.

7 Был произведён анализ результатов прогнозирования контуров объектов моделью SAM. Были выявлены основные достоинства и недостатки модели. Был произведён анализ корреляции значений двух наиболее информативных метрик (полнота и Human Rate). Был произведён анализ изменения метрик при переходе от простых к сложным контурам. Было произведено сравнение значений метрик, полученных для модели SAM, со значениями метрик, полученных для других моделей.

8 Был произведён анализ результатов прогнозирования контуров объектов моделью SAM 2. Были выявлены основные достоинства и недостатки модели. Был произведён анализ корреляции значений двух наиболее информативных метрик (полнота и Human Rate). Был произведён анализ изменения метрик при переходе от простых к сложным контурам. Было произведено сравнение значений метрик, полученных для модели SAM 2, со значениями метрик, полученных для других моделей.

9 Был произведён анализ результатов прогнозирования контуров объектов моделью OneFormer. Были выявлены основные достоинства и недостатки модели. Был произведён анализ корреляции значений двух наиболее информативных метрик (полнота и Human Rate). Был произведён анализ изменения метрик при переходе от простых к сложным контурам. Было произведено сравнение значений метрик, полученных для модели OneFormer, со значениями метрик, полученных для других моделей.

В результате тестирования алгоритмов распознавания контуров объектов, рассмотренных в первой главе, было выявлено следующее:

1 Модель YOLO показала наиболее приемлемый по соотношению качества прогнозирования к скорости работы результат при прогнозировании простых контуров. Однако при прогнозировании средних и сложных контуров модель YOLO справляется значительно хуже остальных трёх моделей.

2 Модель SAM показала отличный результат прогнозирования на изображениях со всеми уровнями сложности контуров. Однако SAM является наиболее медленной моделью. На изображениях с простыми и средними контурами модель OneFormer является более приемлемой по соотношению качества прогнозирования к скорости работы.

3 Модель SAM 2 распознаёт контуры объектов значительно хуже модели SAM. Из достоинств стоит отметить повышение скорости работы в полтора раза по сравнению с прошлой версией модели. Однако модель SAM 2 проигрывает

модели OneFormer как в качестве прогнозирования, так и в скорости работы на изображениях со всеми уровнями сложности контуров, поэтому при построении приложения модель SAM 2 использоваться не будет.

4 Модель OneFormer также показала отличный результат прогнозирования на изображениях со всеми уровнями сложности контуров. Она является наиболее универсальной моделью прогнозирования из всех четырёх рассмотренных моделей. Она лишь немного уступает по соотношению качества прогнозирования к скорости работы модели YOLO на изображениях с простыми контурами и модели SAM на изображениях со сложными контурами.

В **третьей главе** работы были осуществлены проектирование и разработка приложения для гибкой работы с моделями распознавания контуров объектов. Основные этапы процесса проектирования и разработки приведены ниже:

1 Для реализации приложения был выбран фреймворк Django.

2 В качестве базы данных была выбрана PostgreSQL.

3 Была разработана концептуальная модель приложения. Были определены аудитория и функциональность приложение. Также приложение было разделено на два логических блока.

4 Был разработан путь пользователя в приложении. Таким образом схематично был определён порядок взаимодействия между компонентами приложения.

5 Была разработана функциональная модель приложения. На диаграмме UML схематично определены пользователи приложения и их возможности.

6 Была разработана модель данных приложения. Были определены классы для взаимодействия с базой данных, таблицы в базе данных и их поля.

7 Был разработан пользовательский интерфейс, наглядно представлены основные компоненты приложения.

8 Был разработана классов для взаимодействия с пользователем. На ней схематично представлены основные классы и взаимосвязи между ними.

9 Была разработана диаграмма активности, схематично представлен основной вариант использования приложения.

10 Была разработана диаграмма компонентов приложения, в ней все ресурсы приложения были разбиты на три компоненты.

11 Была приведена структура файлов реализованного приложения.

12 Было произведено тестирование основных функции приложения, все тесты были успешно пройдены.

Таким образом, в данной главе было создано решение, позволяющее пользователю гибко выбирать от одной до трёх моделей для прогнозирования изображения. Задача, поставленная в первой и второй главах, была выполнена в полном объёме.

ЗАКЛЮЧЕНИЕ

Основные научные результаты, полученные в диссертационной работе, приведены ниже:

1 Подробно рассмотрен алгоритм распознавания контуров объектов для каждой из версий моделей семейства R-CNN, YOLO, SAM, SAM 2, OneFormer. Тем самым создана обширная теоретическая база для дальнейшего использования моделей.

2 Произведено детальное тестирование алгоритма работы каждой из выбранных моделей. На основе полученных результатов определены сильные и слабые стороны каждого алгоритма, произведено сравнение моделей на изображениях с различными уровнями сложности контуров объектов и различным количеством объектов. Определены наиболее подходящие случаи использования каждой из моделей.

3 Было разработано решение, подходящее для гибкого использования различных моделей в зависимости от нужд пользователя. Пользователю приведена информация о сильных сторонах каждой модели, а также приблизительное время работы каждой из моделей.

Рекомендации по практическому использованию результатов приведены ниже:

1 Полученные теоретические сведения могут быть использованы:

– при разработке обучающих программ и курсов по компьютерному зрению;

– в научных исследованиях и НИОКР-проектах, где требуется выбор оптимальной архитектуры нейросетевой модели для конкретной задачи анализа изображений;

– в инженерных командах, занимающихся интеграцией технологий машинного зрения в промышленные, медицинские, сельскохозяйственные или транспортные системы в качестве справочного материала по выбору подходящего алгоритма;

– при создании методических пособий и документации, регламентирующих использование моделей распознавания контуров объектов в конкретных производственных процессах.

2 Результаты тестирования моделей могут быть использованы:

– для выбора модели под конкретные типы задач;

– для настройки гибридных систем, где выбор модели осуществляется автоматически на основе характеристик изображения;

– в качестве бенчмарка при последующих исследованиях и оптимизации архитектур моделей;

– при оценке производительности вычислительных систем (GPU/CPU) и при планировании инфраструктуры.

3 Разработанное гибкое решение может быть использовано:

– при интеграции в прикладные системы компьютерного зрения, где выбор алгоритма может меняться в зависимости от требований пользователя (точность, скорость, ресурсная нагрузка);

– в автоматизированных системах анализа изображений, где требуется адаптация к условиям съёмки или типам объектов;

– в образовательных и демонстрационных платформах, где важно наглядно показывать различия между архитектурами нейросетевых моделей;

– в инструментах для исследователей и разработчиков, позволяющих ускорить процесс экспериментов с различными моделями;

– в качестве коммерческого программного продукта.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1-А. Харкевич, А.П. Analysis of the basic neural network approaches to the problem of the instance segmentation / Н.Л. Боброва, А.П. Харкевич, В.Ю.Стецко // Актуальные вопросы современной науки и образования: материалы 33-ей международной научно-практической конференции. – Пенза, 2023. – с. 30-37.

2-А. Харкевич, А.П. Overview of the SAM neural network model applicable to the problem of the instance segmentation / Н.Л. Боброва, А.П. Харкевич, В.Ю.Стецко // Актуальные научные исследования: материалы 18-ой международной научно-практической конференции. – Пенза, 2024. – с. 59-62.

3-А. Харкевич, А.П. Overview of the SAM 2 neural network model applicable to the problem of the instance segmentation in video / Н.Л. Боброва, А.П. Харкевич, В.Ю.Стецко // Современное общество, наука и образование: актуальные вопросы, достижения и инновации: материалы 2-ой международной научно-практической конференции. – Пенза, 2025. – с. 27-30.

4-А. Харкевич, А.П. Overview of the OneFormer neural network model applicable to the problem of the instance segmentation in an image / Н.Л. Боброва, А.П. Харкевич, В.Ю.Стецко // Актуальные вопросы современных научных исследований: материалы 17-ой международной научно-практической конференции. – Пенза, 2025. – с. 29-32