

А. В. Андрющенко^{1✉}, Д. А. Гура^{1,2}

Сравнительный анализ нейросетевых архитектур для задач топографического дешифрирования

¹Кубанский государственный технологический университет,
г. Краснодар, Российская Федерация

²Кубанский государственный аграрный университет, г. Краснодар, Российская Федерация
e-mail: antonandr72@gmail.com

Аннотация. В работе представлен сравнительный анализ архитектур искусственных нейронных сетей для задач автоматического топографического дешифрирования данных дистанционного зондирования Земли. Проблема классической ручной оцифровки заключается в высоких временных затратах и неизбежном влиянии человеческого фактора. Цель работы - определение наиболее перспективной модели для автоматизации данного процесса. Рассматриваются три популярные архитектуры: YOLOv8, Mask R-CNN и Segment Anything Model. Сравнение проводится по ключевым критериям: точность выделения границ, скорость обработки, потребность в обучающих данных и требования к аппаратному обеспечению. Установлено, что Segment Anything Model отличается высокой точностью, но требует серьезных вычислительных мощностей и не понимает семантику в автоматическом режиме. Mask R-CNN генерирует качественные контуры, однако существенно уступает конкурентам в скорости. В результате исследования сделан вывод, что оптимальной моделью для поставленных задач является YOLOv8. Данная нейросеть демонстрирует отличные показатели скорости инференса и имеет приоритет для проведения дальнейших эмпирических исследований.

Ключевые слова: топографическое дешифрирование, нейронные сети, YOLOv8, SAM, Mask R-CNN, геоинформационные системы, автоматизация, метрики оценки

A. V. Andryushchenko^{1✉}, D. A. Gura^{1,2}

Comparative analysis of neural network architectures for topographic interpretation tasks

¹Kuban State Technological University, Krasnodar, Russian Federation

²Kuban State Agrarian University, Krasnodar, Russian Federation
e-mail: antonandr72@gmail.com

Abstract. This article presents a comparative analysis of artificial neural network architectures for the tasks of automatic topographic interpretation of Earth remote sensing data. The problem of classical manual digitization lies in high time consumption and the inevitable influence of the human factor. The aim of the study is to determine the most promising model for automating this process. Three popular architectures are considered: YOLOv8, Mask R-CNN, and Segment Anything Model. The comparison is conducted based on key criteria: boundary extraction accuracy, processing speed, training data requirements, and hardware requirements. It is established that Segment Anything Model is characterized by high accuracy but requires significant computing power and lacks semantic understanding in automatic mode. Mask R-CNN generates high-quality contours, but lags significantly behind competitors in speed. As a result of the study, it is concluded that YOLOv8 is the optimal model for the tasks at hand. This neural network demonstrates excellent inference speed and is prioritized for further empirical testing.

Keywords: topographic interpretation, neural networks, YOLOv8, SAM, Mask R-CNN, geographic information systems, automation, and evaluation metrics

Введение

Современные технологии дистанционного зондирования Земли (далее – ДЗЗ) производят большой объем геопространственных данных. Классическая ручная оцифровка данных ДЗЗ требует значительных временных и трудовых затрат. Такая оцифровка предполагает выделение контуров объектов вручную, что неизбежно приводит к ошибкам, ввиду чрезмерного влияния человеческого фактора [1].

Проблему низкой производительности труда и излишнего влияния человеческого фактора возможно решить за счет внедрения в производственный процесс алгоритмов автоматической оцифровки на основе нейронных сетей. При успешном внедрении специалист выполняет функции экспертного контроля, избегая рутинных процессов [2–5].

На данном этапе целью является определение наиболее перспективного вектора исследования, а именно наиболее перспективной модели искусственной нейронной сети для задач автоматической векторизации данных ДЗЗ.

Методы и материалы

В работе рассматриваются три архитектуры: YOLOv8 (YOLOv8-seg), Mask R-CNN и Segment Anything Model (далее – SAM). Выбор обусловлен их пригодностью для обработки данных ДЗЗ.

YOLOv8 – это одностадийная архитектура, которая изначально создавалась с упором на скорость работы в реальном времени. Данная модель показывает высокие результаты по скорости инференса. YOLOv8 способна быстрее конкурентов справляться с сегментацией больших объемов данных. Однако, из-за архитектурных особенностей модель имеет тенденцию к «сглаживанию» углов сложных многоугольников.

Mask R-CNN – это классическая двухстадийная модель для сегментации. Данная модель сначала ищет регионы с потенциальными объектами, а затем генерирует для них маски. Mask R-CNN выдает более качественные и точные контуры по сравнению с YOLOv8, а также отлично справляется с плотной застройкой и наложением объектов друг на друга. Однако, скорость обработки растров заметно ниже конкурентов. Архитектура сложнее в настройке гиперпараметров и более чувствительна к качеству и сбалансированности обучающей выборки, в отличие от аналогов.

SAM – это модель, предварительно обученная на значительном объеме данных. Данная модель обладает высокой точностью выделения контуров. Способна выделять объекты без предварительного дообучения. Однако, в режиме полностью автоматической генерации масок модель не понимает семантику. SAM выделяет на ортофотоплане все видимые объекты без понимания необходимости их выделения (например, автомобили). Для корректной работы в автоматическом режиме SAM возможно комбинировать с другими моделями. Например, используя быструю модель YOLOv8 определить ограничивающую рамку

здания, затем используя модель SAM внутри этой рамки сформировать маску. Также стоит отметить, что SAM, в отличие от других моделей требует больших вычислительных мощностей.

Результаты и обсуждение

Для объективной оценки архитектур YOLOv8, Mask R-CNN и SAM на предмет автоматической оцифровки данных ДЗЗ необходимо определить критерии их оценки.

Поскольку предполагается, что автоматическая оцифровка данных ДЗЗ может использоваться в кадастре, то критичным параметром является точность выделения границ. Это обусловлено тем, что значительная погрешность в определении границ может привести к земельным спорам между собственниками земельных участков и т. п. [2, 6].

Также предполагается, что основными данными ДЗЗ, которые подлежат оцифровке являются ортофотопланы, которые являются массивными растрами с высоким пространственным разрешением. Исходя из этого, выделяется скорость обработки, как один из критериев оценки моделей нейронной сети. Способность модели быстро обрабатывать большие данные является основой экономической эффективности предлагаемых методов автоматической оцифровки [7].

Для корректной работы большинства моделей на конкретной местности необходим набор данных (датасет). Исходя из этого формируется еще один критерий – потребность в данных. Он определяется объемом датасета, необходимого для того, чтобы модель начала уверенно распознавать специфические объекты. Например, сложные конфигурации крыш, заборы, грунтовые дороги и т. п. [8].

Не менее важный критерий – это требование к аппаратному обеспечению. Под этим критерием понимается объем видеопамати и вычислительная мощность графического процессора, необходимые для обучения и применения модели искусственной нейронной сети.

После проведения анализа критериев применительно к каждой рассматриваемой модели искусственной нейронной сети была сформирована таблица соответствия критериев и моделей искусственной нейронной сети (табл. 1). Соответствие критериям определялось в соответствии со шкалой оценки: «Очень низко» – «Низко» – «Средне» – «Высоко» – «Очень высоко».

Таблица 1. Сравнительный анализ архитектур искусственной нейронной сети

Критерий	YOLOv8	Mask R-CNN	SAM
Точность границ	Средняя	Высокая	Очень высокая
Скорость работы	Очень высокая	Средняя	Низкая
Потребность в данных	Высокая	Высокая	Очень низкая
Требования к аппаратуре	Низкие	Средние	Высокие

Исходя из данных в таблице 1, можно сделать вывод о том, что наилучшей моделью искусственной нейронной сети для задачи автоматического дешифрирования данных ДЗЗ является YOLOv8. Данная модель имеет приоритет для дальнейших эмпирических тестов и будет протестирована в первую очередь во время будущих исследований.

Заключение

В ходе проведенного исследования был выполнен сравнительный анализ архитектур YOLOv8, Mask R-CNN и SAM. Целью работы являлось определение наиболее перспективной модели искусственной нейронной сети для задач автоматической векторизации данных ДЗЗ.

Оценка нейросетевых моделей проводилась на основе комплекса критериев, включая точность выделения границ, скорость обработки, потребность в обучающих данных и требования к аппаратному обеспечению.

Модель SAM обладает высокой точностью выделения контуров. Однако, в режиме автоматической генерации масок она не понимает семантику и выделяет все видимые объекты. Также данная архитектура требует больших вычислительных мощностей в отличие от других моделей.

Двухстадийная архитектура Mask R-CNN выдает более качественные и точные контуры по сравнению с YOLOv8 и отлично справляется с плотной застройкой. При этом скорость обработки растров у нее заметно ниже конкурентов, а сама модель сложнее в настройке гиперпараметров и чувствительнее к обучающей выборке.

Одностадийная архитектура YOLOv8 демонстрирует высокие результаты по скорости инференса. Данная модель способна быстрее конкурентов выполнять сегментацию больших объемов данных.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Тематическое дешифрирование аэроснимков лесных территорий на основе концептуального моделирования / К. В. Шошина, Р. А. Алешко, В. В. Березовский [и др.] // Онтология проектирования. – 2023. – Т. 13, № 3(49). – С. 437–454.
2. К вопросу о применении нейросетевого алгоритма для мониторинга состояния городской среды / Р. А. Дьяченко, Д. А. Гура, Д. А. Беспятчук [и др.] // Экологические системы и приборы. – 2023. – № 8. – С. 42–50. – DOI 10.25791/esip.8.2023.1392. – EDN KСЕСРТ.
3. Разработка алгоритма семантической сегментации аэрофотоснимков реального времени / Ю. Б. Блохинов, В. А. Горбачев, Ю. О. Ракутин, А. Д. Никитин // Компьютерная оптика. – 2018. – Т. 42, № 1. – С. 141–148.
4. Семантическая сегментация данных дистанционного зондирования Земли при помощи нейросетевых алгоритмов / А. А. Друки, В. Г. Спицын, Ю. А. Болотова, А. А. Башлыков // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2018. – Т. 329, № 1. – С. 59–68.
5. Семантическая сегментация спутниковых снимков аэропортов с помощью свёрточных нейронных сетей / В. А. Горбачев, И. А. Криворотов, А. О. Маркелов, Е. В. Котлярова // Компьютерная оптика. – 2020. – Т. 44, № 4. – С. 636–645.
6. Береговой, Д. В. Методика автоматизированного создания топографического плана на основе съёмки с беспилотного летательного аппарата / Д. В. Береговой, М. Г. Мустафин // Геодезия и картография. – 2018. – Т. 79, № 9. – С. 30–36.

7. Алгоритмы семантической сегментации разновременных многоспектральных изображений с использованием глубокого обучения / Ю. Ю. Громов, Ю. Г. Веселов, И. Н. Ищук, В. В. Родионов // Приборы и системы. Управление, контроль, диагностика. – 2025. – № 2. – С. 1–16. – DOI 10.25791/pribor.2.2025.1555.

8. Дьяченко, Р. А. К вопросу об увеличении производительности машинного обучения на этапе выборки данных при решении задач классификации / Р. А. Дьяченко, П. А. Косолапов, Д. А. Гура // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2022. – № 4. – С. 146–155. – DOI 10.17308/sait/1995-5499/2022/4/146-155.

© А. В. Андриященко, Д. А. Гура, 2026