

Емельянов Антон Владимирович

**Дешифрирование и векторизация аэро- и космофотоснимков
методами машинного обучения для обновления
геопространственной информации**

Специальность 1.2.1.

Искусственный интеллект и машинное обучение

Автореферат

диссертации на соискание учёной степени
кандидата технических наук

Работа выполнена в Федеральном государственном автономном образовательном учреждении высшего образования "Московский физико-технический институт (национальный исследовательский университет)".

Научный руководитель: кандидат технических наук
Князь Владимир Александрович

Ведущая организация: Федеральное государственное бюджетное учреждение науки Институт проблем управления им. В. А. Трапезникова Российской академии наук

Защита состоится 26 декабря 2025 г. в 16:00 часов на заседании диссертационного совета ФПМИ.1.2.1.015, созданного на базе федерального государственного автономного образовательного учреждения высшего образования "Московский физико-технический институт (национальный исследовательский университет)" по адресу: 141701, Московская область, г. Долгопрудный, Институтский переулок, д. 9.

С диссертацией можно ознакомиться в библиотеке МФТИ, Физтех и на сайте организации <https://mipt.ru>.

Автореферат разослан 26 октября 2025 года.

Ученый секретарь
диссертационного совета
ФПМИ.1.2.1.015,
канд. техн. наук

Войтиков Константин Юрьевич

Общая характеристика работы

Актуальность темы.

Число задач, в которых требуется актуальное и достоверное геоинформационное обеспечение расширяется, увеличивается с каждым годом. Основным источником геоинформационного обеспечения в настоящее время является аэро- и космофотосъёмка. Однако сложность первичного дешифрирования аэро- и космических снимков вызывает значительное количество ошибок и создаёт временную задержку в доставке актуальных векторных карт потребителям геоинформационных данных.

С другой стороны, в последнее десятилетие произошёл технологический и научный прорыв в методах обработки изображений, который связан с появлением высокопроизводительных массово-параллельных вычислителей и средствами эффективного обучения глубоких свёрточных нейронных сетей. Несмотря на значительный прогресс нейросетевых методов, существующие на текущий момент методы не обеспечивают качества обработки спутниковых и аэрофотоснимков, сравнимого с качеством построения векторных карт человеком-оператором.

Данный факт обуславливает актуальность и создаёт предпосылки для разработки методов и алгоритмов интеллектуального анализа аэро- и космоснимков для обновления и векторизации геопространственной информации в обеспечение создания отечественного программного комплекса обновления картографического обеспечения.

Степень разработанности темы.

Развитие методов семантической сегментации и векторизации аэро- и космофотоснимков, берущее начало на заре компьютерного зрения, прошло последовательную эволюцию от аналитических методов к глубоким свёрточным сетям, где ключевую роль сыграло наращивание вычислительных возможностей. Аналитические методы сегментации аэрофотоснимков включают пороговые алгоритмы (Оцу, адаптивное пороговое разделение), операторы выделения краев (Собея, Кэнни) и морфологические операции (эрозия, дилатация). Методы векторизации основаны на алгоритмах аппроксимации контуров

(Рамера-Дугласа-Пекера), скелетизации (метод медиальной оси, алгоритм Чжана-Суэня) и регуляризации через минимизацию энергетических функционалов. Для анализа текстур применяются статистики Харалика и матрицы смежности, позволяющие выделять однородные зоны по спектральным и пространственным признакам.

Основу современных исследований составляют работы в области семантической сегментации, где фундаментальный вклад внесли Дж. Лонг с разработкой полностью сверточных сетей (FCN), О. Роннебергер с архитектурой U-Net, специально оптимизированной для биомедицинских и спутниковых изображений, а также Л. Чэнь с серией моделей DeepLab. Параллельно развивается направление детектирования объектов на снимках, где значимыми являются работы Р. Гиршика по R-CNN, Дж. Редмона по YOLO и В. Лю по SSD-архитектурам. В области прямого прогнозирования векторных данных перспективные разработки связаны с применением графовых нейронных сетей и методов обучения с подкреплением, где можно отметить исследование Я. Лина по распознаванию инженерных чертежей. Для решения специализированных задач геоинформатики сохраняют актуальность классические алгоритмы машинного обучения, такие как случайный лес и метод k-ближайших соседей, теоретические основы которых заложены в работах А.С. Потапова по искусственному интеллекту и Д. Черниховского по применению k-NN для оценки лесных массивов. Данные алгоритмы позволили достигнуть высокого качества решения задач.

Целью данной работы является разработка методов по повышению качества дешифрирования и векторизации аэрокосмических изображений земной поверхности на основе машинного обучения в области разработки интеллектуальных методов анализа и обновления геопространственной информации.

Для достижения поставленной цели необходимо было решить следующие **задачи**:

1. Провести анализ существующих современных методов сегментации и векторизации аэро- и космифотоснимков.

2. Разработать метод точного воссоздания полигональных границ зданий для создания векторизованной маски для использования в различных приложениях.
3. Разработать методику обучения нейросетевых алгоритмов векторизации аэро- и космоснимков на основе технологии обучения с подкреплением.
4. Разработать метод сегментации аэрофотоснимков с применением теории графов для использования существующих связей между объектами и областями на снимках.
5. Собрать выборку, включающую в себя аэро- и космоснимки из открытых источников, а также полученных в ходе натурных экспериментов, с целью проведения вычислительных испытаний для разработанных методов.

Научная новизна:

1. Впервые разработан метод сегментации и векторизации спутникового снимка для построения иерархического графа семантической разметки изображения, учитывающего не только семантическую роль объектов на изображении, но и связи и взаимоотношения между ними.
2. Впервые разработан алгоритм сегментации и векторизации снимка, имитирующий логику и действия оператора-разметчика и развивающий методы обучения с подкреплением.
3. Создана обширная обучающая выборка, содержащая не только пары аэро- и космофотоснимков и разметки к ним, но и обширную информацию о хронологии и этапах разметки, выполняемых оператором-дешифровщиком, включая масштабы изображений, выбранные для разметки заданного объекта, порядок обхода объектов и другие виды метainформации. На разработанной выборке проведены вычислительные эксперименты для реализованных методов.

Научная и практическая значимость решения обозначенной проблемы состоит в создании методов анализа данных космической и аэрофотосъёмки, позволяющих автоматически и с высоким качеством решать задачи семантической сегментации снимка, выявления изменений в наблюдаемой сцене и получения производной (векторной) информации, необходимой для оперативного обновления геоданных. Разработанные методы позволят, с одной стороны, более эффективно решать задачи обновления геопространственной информации, такие как выявление изменений в ландшафте и инфраструктуре, корректировка и пополнение базы геоданных, а с другой стороны, станут научно-методической основой для решения более сложных задач анализа геопространственной информации, таких как: создание иерархических моделей геопространственной инфраструктуры, интеллектуальное автоматическое оперативное создание карт, автоматическое интеллектуальное планирование сельско-хозяйственных работ, автоматическое интеллектуальное городское планирование, автоматический прогноз потенциальных природных и техногенных происшествий.

Методология и методы исследования. В работе использовались методы цифровой обработки изображений, анализа данных и машинного обучения. Для определения работоспособности реализованных методов и сравнения их с ведущими современными подходами использовались открытые выборки, являющиеся стандартом для тестирования алгоритмов сегментации и векторизации, содержащие реальные данные. Разработка программного кода производилась на языке программирования Python, с использованием фреймворка для глубокого обучения PyTorch, а так же библиотеки OpenCV, scikit-learn и Pillow.

Основные положения, выносимые на защиту:

1. Разработанный метод многоэтапной регуляризации обеспечивает повышение точности извлечения контуров зданий на изображениях дистанционного зондирования Земли по сравнению с современными ведущими подходами на 5-10% по метрикам средней точности и полноты и около 1% по коэффициенту Жаккара.

2. Разработанный метод векторизации аэрофотоснимков на основе обучения с подкреплением обрабатывает изображения быстрее (на 60%) и с меньшими использованиями памяти вычислительных средств (на 21%), чем многоуровневые архитектуры, при незначительном снижении качества, что допустимо для задач векторизации в реальном времени.
3. Разработанная двухэтапная архитектура на основе нейронных сетей и теории графов демонстрирует высокий уровень эффективности при обработке аэрофотоснимков сложной структуры и обеспечивает увеличение точности семантической сегментации за счет комбинации методов глубокого обучения и графовых подходов. По результатам эксперимента прирост коэффициента Жаккара 5-10% по сравнению с современными автоматическими методами.

Достоверность полученных результатов подтверждается проведенными экспериментами для используемых методов, однородностью разметки аэро- и космофотоснимков, составляемой для одного снимка различными независимыми профессиональными дешифровщиками, а также формализованностью структуры векторной карты местности и процесса разметки, позволяющей создать единообразное математическое представление в виде графа сцены.

Апробация работы. Основные результаты по теме диссертации изложены в 6 печатных изданиях, индексируемых Web of Science и Scopus, 1 из которых изданы в журналах списка K1 и входят в собственный перечень МФТИ. Основные результаты работы докладывались на следующих научных конференциях:

1. ISPRS TC I Mid-term Symposium on Intelligent Sensing and Remote Sensing Application, Changsha, 2024;
2. ISPRS TC III Mid-term Symposium “Beyond the canopy: technologies and applications of remote sensing”, Belém, 2024;
3. International Workshop on “Photogrammetric Data Analysis” – PDA24, Moscow, 2024;

4. International workshop on Photogrammetric techniques for environmental and infraStructure monitoring, Biometry and Biomedicine – PSBB25, Moscow, 2025.

Также результаты диссертационного исследования использованы при реализации проекта российского научного фонда №24-21-00269 ”Дешифрование и векторизация аэро- и космофотоснимков методами машинного обучения для обновления геопространственной информации”, научно-исследовательских работ, таких как проект ”САПФИР” для разработки программного обеспечения, использующего высокопроизводительные вычислительные средства, выполняемый в рамках субсидии Минпромторга России, и проект ”СИНТЕЗ” для разработки технологии повышения реалистичности тепловизионных изображений, формируемых на основе высокодетализированных трехмерных моделей местности. В первом случае для семантической сегментации использовался разработанный метод на основе визуального трансформера и графовой нейронной сети, во втором случае - разработанные методы векторизации для улучшения качества построения трехмерных моделей местности.

Личный вклад. Все основные результаты получены автором лично. Также автор участвовал в постановке исследовательских задач, выборе методов их решения, анализе экспериментальных данных, интерпретации и обсуждении результатов, и формулировке выводов работы. Во всех выносимых на защиту результатах личный вклад диссертанта является основным и определяющим.

Содержание работы

Во **введении** обосновывается актуальность исследований, проводимых в рамках данной диссертационной работы, формулируется цель, ставятся задачи работы, излагается научная новизна и практическая значимость представляемой работы.

Первая глава содержит обзор литературы, существующих подходов к задачам сегментации и векторизации аэро- и космофотоснимков, задаче извлечения границ зданий с изображений земной поверхности, а также распространенные открытые базы данных.

Анализ современных методов обработки аэрофотоснимков показывает эволюцию от традиционных подходов компьютерного зрения к глубоким нейросетевым архитектурам. Ключевой тенденцией является переход от семантической сегментации к комплексным решениям, сочетающим CNN, трансформеры и графовые нейронные сети. Особое внимание уделяется решению проблемы дисбаланса классов и улучшению качества векторизации через регуляризацию границ.

Одной из перспективных направлений признана интеграция обучения с подкреплением (RL) с методами компьютерного зрения, что позволяет формулировать задачу векторизации как процесс последовательного принятия решений. Это обеспечивает адаптивность к сложным пространственным зависимостям и улучшает обработку объектов со сложной структурой. Современные исследования демонстрируют эффективность гибридных подходов, сочетающих преимущества различных архитектур для достижения высокой точности при работе с реальными данными дистанционного зондирования.

Вторая глава посвящена исследованию точного воссоздания полигональной границы при одновременном извлечении векторизованной маски строения в качестве выходных данных для непосредственного использования в различных приложениях. Для этого представлены 2 последовательных алгоритма, каждый из которых автоматически извлекает контуры зданий посредством комбинации методов бинарной семантической сегментации, регуляризации и векторизации.

Сначала был разработан метод автоматизированного извлечения контуров зданий основан на анализе локального контекста с использованием матриц

окрестности. Метод реализует трехэтапный процесс: (1) бинарную семантическую сегментацию с использованием архитектуры UNetFormer, (2) регуляризацию предсказаний на основе анализа матриц окрестности, (3) векторизацию результатов в полигональное представление.

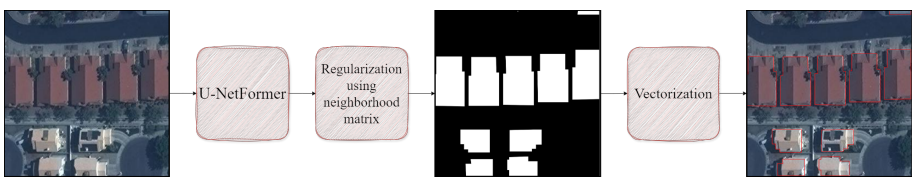


Рисунок 1 — Структурная схема алгоритма матрично-окрестностной регуляризации

Начальный этап метода включает бинарную семантическую сегментацию изображений дистанционного зондирования для идентификации зданий. Для выполнения этой задачи использовалась нейронная сеть UNetFormer.

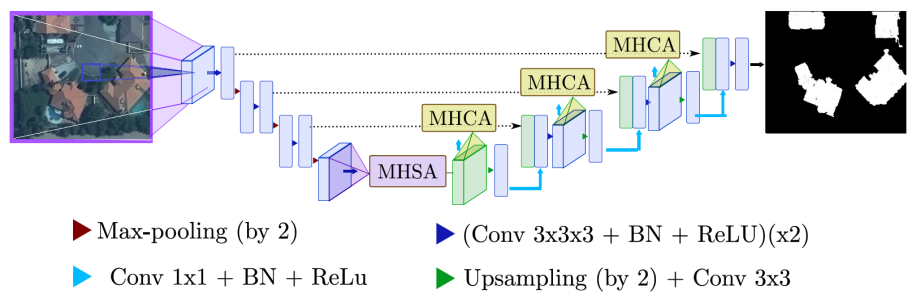


Рисунок 2 — Структура сверточной нейронной сети, используемой для бинарной сегментации.

Регуляризация основана на анализе матриц окрестности размера 3×3 , где для каждого пикселя вычисляется параметр K — сумма элементов матрицы, принадлежащих классу "здание". Критерий коррекции классификации:

- $K \leq 3$ — ошибочно сегментированные пиксели
- $K = 4$ — вершины выпуклых углов
- $5 \leq K \leq 8$ — граничные точки

– $K = 9$ — внутренние точки

Также отметим, что на первом этапе алгоритма точки со значением $K = 8$ добавляются в класс ”здание если нейронная сеть не распознала их как представителей нужного класса.

Метод обеспечивает эффективное устранение шумов и артефактов сегментации, улучшая качество выделения контуров зданий за счет учета локальных пространственных закономерностей.

В процессе доработки предыдущего алгоритма была создана методология, которая реализует многоэтапную цепочку обработки, которая последовательно выполняет: (1) бинарную семантическую сегментацию с использованием глубокой сверточной нейронной сети, (2) комплексное уточнение границ с помощью интегрированных методов регуляризации и (3) геометрическую векторизацию улучшенных выходных данных. Структура алгоритма, интегрирующая эти методы регуляризации, проиллюстрирована на Рисунке 3.

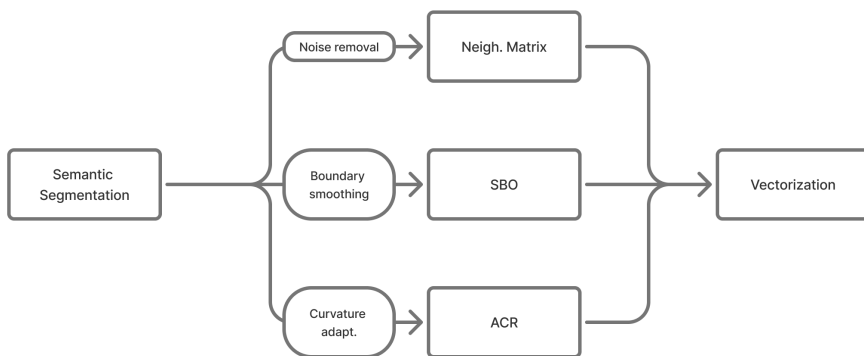


Рисунок 3 — Структура предлагаемого алгоритма, объединяющая регуляризацию матрицы соседства, оптимизацию спектральной границы и адаптивное уточнение контура.

Разработанная методология использует сложный подход глубокого обучения для обнаружения контуров зданий из изображений дистанционного

зондирования, начиная с бинарной сегментации с использованием улучшенной архитектуры UNetFormer. Эта продвинутая нейронная сеть расширяет возможности традиционных фреймворков U-Net за счёт интеграции мощных механизмов внимания на основе трансформеров, обеспечивая более комплексный анализ пространственных отношений на множестве масштабов.

Многоэтапный процесс регуляризации:

1. Начальная регуляризация на основе матрицы окрестности

Первый этап конвейера регуляризации использует подход локальной пространственной согласованности для исправления ошибок классификации на уровне пикселей. Для каждого пикселя $p_{i,j}$, классифицированного как "здание" (метка класса 1), мы строим матрицу окрестности 3×3 $N(p_{i,j})$, охватывающую центральный пиксель и его восемь непосредственных соседей. Эта матрица служит окном локального контекста, позволяя:

- **Обнаруживать неправильно классифицированные пиксели:** Пиксель помечается как потенциально неправильно классифицированный, если его метка противоречит большинству его соседей (например, изолированный пиксель "здания" окруженный пикселями "не-здания").
- **Применять вероятностную коррекцию:** Конечная метка $p_{i,j}$ переопределяется на основе взвешенного голоса его соседей, где веса обратно пропорциональны их евклидовому расстоянию от $p_{i,j}$. Этот шаг эффективно удаляет шум "соли и перца" (случайно возникающие черные и белые пиксели), сохраняя законные структуры малого масштаба.

2. Спектральная оптимизация границ (SBO) для глобальной согласованности

Для обеспечения топологической согласованности между сегментами зданий мы моделируем выход сегментации как ненаправленный граф

$G = (V, E)$, где Узлы V соответствуют пикселям, помеченным как ”здание Рёбра E соединяют пространственно смежные пиксели (8 соседей), с весами w_{ij} , определенными гауссовым ядром сходства:

$$w_{ij} = \exp \left(-\frac{|p_i - p_j|_2^2}{2\sigma_d^2} - \frac{|I_i - I_j|_2^2}{2\sigma_I^2} \right), \quad (1)$$

где σ_d и σ_I контролируют чувствительность к пространственному расстоянию и вариации интенсивности соответственно.

Лапласиан графа $L = D - W$ (где D — степень матрицы, а W — взвешенная матрица смежности) кодирует глобальную структуру областей зданий. Мы решаем задачу спектральной оптимизации:

$$\min_f (f^T L f + \lambda |f - y|^2), \quad (2)$$

где f — регуляризованное поле меток, y — начальная сегментация, и λ — параметр компромисса. Этот шаг устраняет фрагментированные области и сглаживает нерегулярные границы, уважая края изображения.

3. Адаптивное уточнение контура (ACR) для геометрической точности

Финальный этап уточняет границы зданий через явную адаптацию кривизны. Для каждого извлеченного контура границы C :

- **Вычисляем локальную кривизну** κ_i в каждой точке i с помощью производно-ориентированного оценщика:

$$\kappa_i = \frac{|x'iy''i - y'ix''i|}{(x_i'^2 + y_i'^2)^{3/2}}, \quad (3)$$

где производные аппроксимируются через центральные конечные разности (например, $x' = \frac{xi+1-xi-1}{2}$).

- **Регулируем силу сглаживания** с помощью кривизно-адаптивного ядра:

$$h_i = \exp \left(-\frac{\kappa_i^2}{2\sigma_c^2} \right), \quad \sigma_c = \text{толерантность кривизны}. \quad (4)$$

Области с высокой кривизной (например, углы с $\kappa_i > \sigma_c$) подвергаются слабому сглаживанию, сохраняя резкие черты, в то

время как сегменты с низкой кривизной (прямые края) агрессивно регуляризуются.

- **Передискретизируем контур** используя B-сплайн интерполяцию для обеспечения равномерного расстояния между точками, критически важного для высококачественной векторизации. Конечный контур реконструируется как:

$$\mathcal{C}_{\text{refined}} = \sum ih_i \cdot K(|\mathcal{C}_i - \mathcal{C}|) * \mathcal{C}, \quad (5)$$

где K — гауссово ядро сглаживания.

В обоих методах нейронные сети оптимизированы с использованием функции бинарной кросс-энтропии:

$$L = -\frac{1}{N} \sum_{i=1}^N [x_i \log(\sigma(y_i)) + (1 - x_i) \log(1 - \sigma(y_i))] \quad (6)$$

где $\sigma(x) = \frac{1}{1+e^{-x}}$ — сигмоидная функция активации.

Разработанные подходы прошли экспериментальную проверку на наборе данных CrowdAI Mapping Challenge, показав превосходство над существующими методами по ключевым метрикам качества, которыми являлись IoU, Average Precision (AP) и Average Recall (AR), которые вычислялись с помощью точности и полноты:

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

Для оценки производительности в Таблице 2 представлены количественные результаты предложенных методов вместе с сравнительными эталонными показателями современных подходов. Этот сравнительный анализ позволяет оценить относительную производительность алгоритмов по сравнению с существующими решениями на сопоставимых данных.

Третья глава посвящена разработке нового подхода к векторизации аэрофотоснимков, основанного на обучении с подкреплением (RL). Основная цель

Метод	AP	AR	IoU
Mask R-CNN	41.9	47.6	-
PolyMapper	55.7	62.1	-
PolyWorld	63.3	75.4	91.3
Матрица окрестности	64.1	75.1	91.2
Многоэтапная регуляризация	69.8	80.3	92.0

Таблица 1 — Результаты на тестовом наборе данных CrowdAI для всех экспериментов по извлечению и полигонизации зданий.

— разработать адаптивную модель, способную принимать последовательные решения при выделении объектов, тем самым повышая точность векторизации по сравнению с традиционными методами. Ключевое отличие предлагаемого подхода заключается в применении RL к задачам векторизации, что позволяет модели учитывать пространственный контекст и динамически адаптироваться к различным типам объектов.

Основу методологии составляет гибридная архитектура, интегрирующая три ключевых компонента:

- **Модуль извлечения признаков** на базе модифицированной сети ResNet-50 с механизмами пространственного внимания и многоуровневым пулингом (ASPP) для обработки многомасштабного контекста
- **Модуль последовательного контекста** на основе двунаправленных GRU-сетей для анализа пространственных зависимостей
- **Агент обучения с подкреплением** по архитектуре Dueling Double DQN для оптимизации процесса векторизации

Математическая модель агента RL описывается уравнением:

$$Q(s,a) = V(s) + \left(A(s,a) - \frac{1}{|A|} \sum A(s,a') \right) \quad (9)$$

Оптимизация обучения реализована через:

- Модифицированный алгоритм DQN с механизмом Double Q-learning
- Приоритизированное воспроизведение опыта
- Параметризованную стратегию исследования Noisy Nets

Система вознаграждения включает многокомпонентную функцию:

$$R(s,a) = \alpha R_{\text{geo}} + \beta R_{\text{topo}} + \gamma R_{\text{eff}} \quad (10)$$

с коэффициентами $\alpha = 0.6$, $\beta = 0.3$, $\gamma = 0.1$.

Для обучения используется композитная функция потерь:

$$L = 0.7 \cdot \text{Dice} + 0.3 \cdot \text{Focal}(\gamma = 2, \alpha = 0.8) \quad (11)$$

Оптимизация проводится с использованием алгоритма AdamW со скоростью обучения 3×10^{-4} и весовым затуханием 1×10^{-5} .

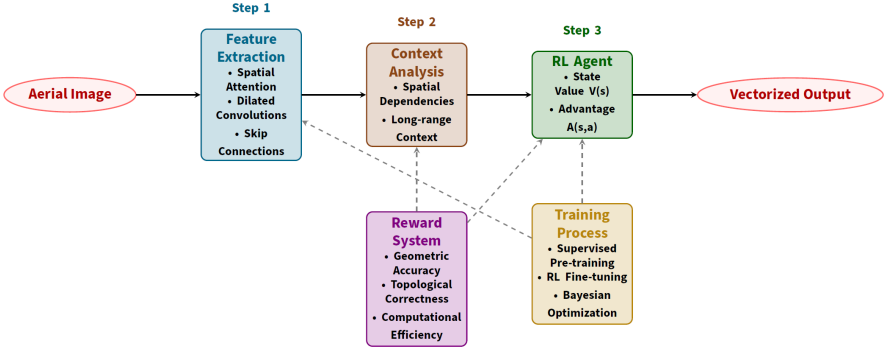


Рисунок 4 — Общая схема предложенного алгоритма

Разработанный подход прошел комплексную экспериментальную проверку на наборе данных CrowdAI Mapping Challenge с использованием нескольких метрик оценки.

Оценка геометрической точности использует среднее значение Intersection over Union (mIoU) в качестве основной метрики.

Intersection-over-Union (IoU) или индекс Жаккара — это отношение площади пересечения предсказанной и истинной маски к их объединению:

$$IoU = \frac{TP}{TP + FP + FN} \quad (12)$$

В данной работе mIoU вычисляется для всех классов объектов как:

$$mIoU = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FP_c + FN_c} \quad (13)$$

где C представляет количество классов (здания, дороги, водоемы и т.д.), TP_c обозначает истинно положительные срабатывания для класса c , FP_c указывает ложноположительные срабатывания, а FN_c учитывает ложноотрицательные срабатывания.

Качеству границ уделяется особое внимание через оценку Boundary F1-score (BD-F1), которая вычисляется с использованием значений точности и полноты:

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN} \quad (14)$$

$$BD - F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (15)$$

Оценка топологической корректности включает две специализированные метрики: Junction Accuracy измеряет процент правильно идентифицированных точек пересечения между линейными объектами:

$$J_{acc} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(\|J_{pred}^i - J_{gt}^i\| < \epsilon_{junction}) \quad (16)$$

где $\epsilon_{junction}$ представляет радиус принятия в 5 пикселей, а Cycle Consistency оценивает целостность замыкания полигонов с помощью анализа связности вершин на основе теории графов.

Значения вычисленных метрик приведены в таблице 1. Для понимания уровня эффективности алгоритма в таблицу также включены результаты ведущих методов на аналогичных данных.

Качественная оценка показывает надежную работу системы для различных категорий объектов, причем контуры зданий демонстрируют особенно точное определение границ (средняя ошибка позиционирования вершин 1.2 пикселя), а дорожные сети сохраняют превосходную топологическую связность (Junction Accuracy 93.4% и Cycle Consistency 94.1%). Выходные данные векторизации сохраняют картографическое качество с четко определенными вершинами и

Метод	mIoU	BD-F1	Время (мс)	Память (ГБ)
U-Net	0.75	0.66	120	6.2
Предложенный RL	0.82	0.79	85	9.8
HRNet	0.84	0.81	210	12.4

Таблица 2 — Результаты на тестовом наборе данных CrowdAI для всех экспериментов по векторизации.

минимальным количеством избыточных точек, уменьшая среднюю сложность полигонов на 28% по сравнению с традиционными подходами на основе сегментации.

Количественный анализ подтверждает значительное повышение производительности по сравнению с подходом на основе U-Net. Однако стоит отметить, что разработанный метод немного уступает по эффективности многоуровневым архитектурам, таким как HRNet, но обрабатывает изображение быстрее, что может быть важно для решения задач векторизации в реальном времени. Анализ использования памяти показывает особенно впечатляющие результаты для обработки больших областей, что позволяет проводить практическое развертывание на стандартном оборудовании.

Четвертая глава посвящена разработке нейросетевой модели для извлечения семантической и векторной информации о местности на основе анализа аэрофотоснимка. Разработанная нейросетевая модель на первом этапе использует визуальный трансформер для извлечения глубоких признаков из входного изображения. Затем графовая нейронная сеть выполняет кластеризацию извлеченных глубоких признаков для получения семантической сегментации изображения.

Для обучения и оценки разработанной нейросетевой модели была собрана и аннотирована специальная выборка, содержащая более 10 тысяч аэрофотоснимков, сделанных в разные годы и сезоны и представляющих различные типы объектов.

При создании выборки SVAI учитывались следующие основные требования:

- большой объём данных: 10 000 изображений или более;
- разнообразие данных: аэрофотоснимки должны содержать различные участки земной поверхности (городская застройка, транспортная сеть, сельскохозяйственные угодья и т.п.);
- разнообразие условий: выборка должна содержать данные, полученные в разные сезоны, время суток;
- разнообразие применяемых платформ: снимки должны быть получены с различных носителей и различными сенсорами.

Подготовленная выборка содержит более 12 000 изображений, приведённых к одному размеру (512×512 пикселей) и одному масштабу съёмки (0,34 м/пиксель). Для подготовки выборки использованы материалы спутниковой съёмки (Bing и Google Earth) и съёмки с беспилотных летательных аппаратов (БПЛА). Собранные данные охватывают период 12 лет и включают как появление новых объектов, так и исчезновение существовавших ранее.

В выборке SVAI представлены различные типы зданий и сооружений, автомобильные и железные дороги, различные типы ландшафтов (лесные участки, городская застройка различных типов, водные участки), полученные в разное время.

Для аннотации данных использован формат бинарных масок, традиционно применяемый в задачах обнаружения изменений. На рис. 5 показаны примеры изображений и аннотаций из выборки SVAI. Временной период изменений: середина 2018 года – середина 2024 года.

Для исследований в области векторизации аэрофотоснимков в выборку SVAI включены векторные данные об объектах. Векторные аннотации формировались двумя способами. Первый – ручная разметка оператором-дешифровщиком. Второй – создание векторной разметки с привлечением ресурса «Открытая карта улиц» (OpenStreetMap, OSM).

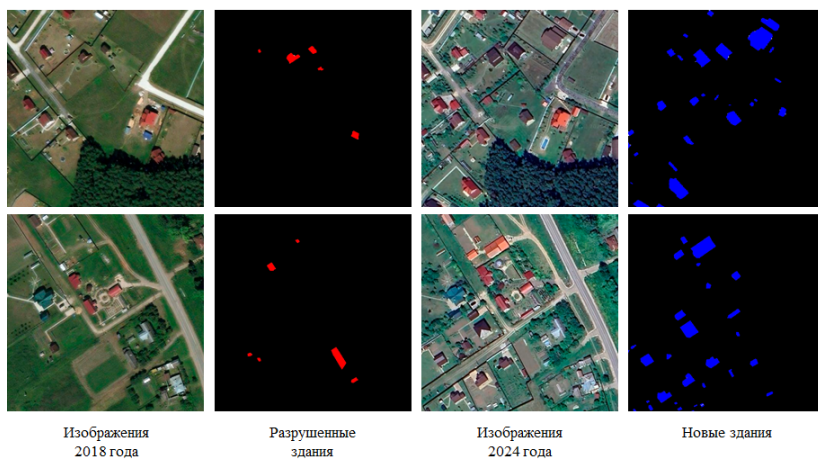


Рисунок 5 — Примеры изображений из набора данных Segmentation and Visualization Aerial Images. Изображения в столбцах 1 и 3 были получены со спутников Bing и Google в 2018 и 2024 годах соответственно; столбцы 2 и 4 показывают изменения (разрушенные объекты показаны красным, а новые появившиеся объекты показаны синим)

Пример данных из выборки, предназначенной для задач векторизации аэрофотоснимков, показан на рис. 6. Для первичной разметки данных использовались данные OpenStreetMap, которые впоследствии проверялись (и корректировались в случае необходимости) оператором-дешифровщиком с формированием журнала действий оператора, необходимым для исследований методов обучения с подкреплением в задачах обнаружения изменений и векторизации аэрофотоснимков.

Разработанная двухэтапная архитектура нейросетевой модели для семантической сегментации аэрофотоснимков основана на принципах работы оператора-дешифровщика. На первом этапе осуществляется анализ всего изображения с выделением областей интереса, на втором – детальное изучение выделенных областей с формированием семантических объектов.

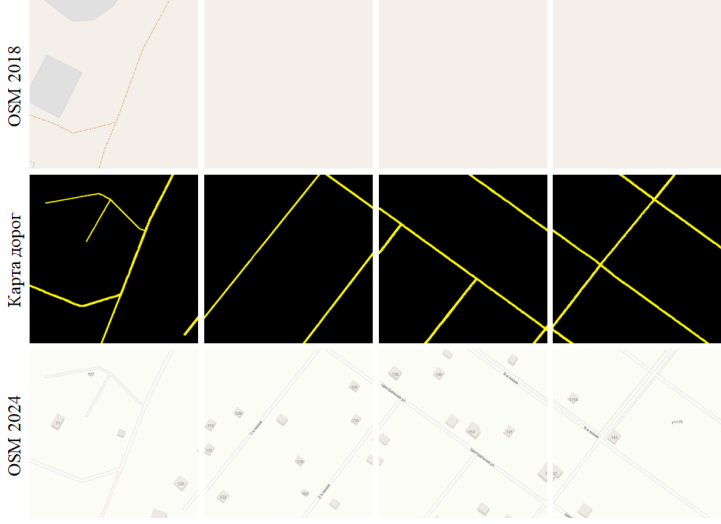


Рисунок 6 — Примеры изображений из выборки SVAI. Верхний и нижний ряды — это изображения OSM за 2018 и 2024 годы соответственно, а средний ряд — это изменения в дорожной карте за этот период времени

Для извлечения признаков используется визуальный трансформер (ViT), обученный по методу самообучения DINO. Модель преобразует входное изображение I размером $h \times w$ в k^2 фрагментов (токенов) размером:

$$s = \frac{h}{k} \times \frac{w}{k} \quad (17)$$

Выходной вектор признаков Z имеет размерность $s \times t$.

Изображение представляется в виде графа $G = \mathcal{V}, \mathcal{E}$, где сходство между областями задается матрицей:

$$W = ZZ^T \in \mathbb{R}^{k \times k} \quad (18)$$

Для предотвращения выделения изолированных узлов используется критерий нормализованного разреза:

$$Ncut(A, B) = \frac{\sum_{u \in A, v \in B} w(u, v)}{\sum_{u \in A, v \in \mathcal{V}} w(u, v)} + \frac{\sum_{u \in A, v \in B} w(u, v)}{\sum_{u \in B, v \in \mathcal{V}} w(u, v)} \quad (19)$$

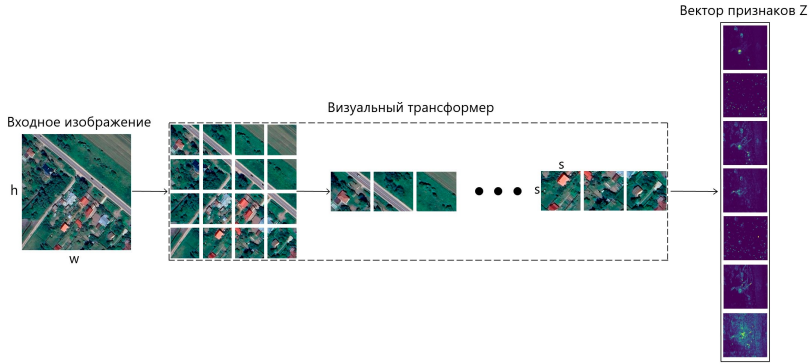


Рисунок 7 — Схема построения графа на основе матрицы подобиия признаков

Кластеризация осуществляется с помощью графовой нейронной сети:

$$\hat{X} = GNN(n, W; \Theta_{GNN}), \quad C = MLP(\hat{X}; \Theta_{MLP}) \quad (20)$$

Функция потерь оптимизируется по модифицированному критерию нормализованного разреза:

$$\mathcal{L}_{Ncut} = -\frac{Tr(C^T W C)}{Tr(C^T D C)} + \left\| \frac{C^T C}{\|C^T C\|_F} - \frac{\mathbb{I}_3}{\sqrt{3}} \right\|_F \quad (21)$$

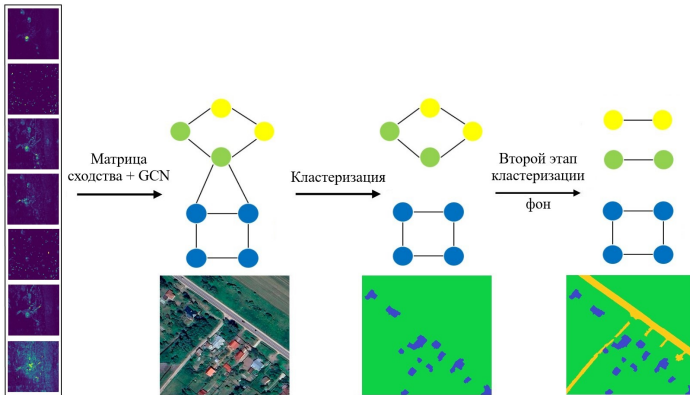


Рисунок 8 — Двухэтапная процедура кластеризации

Оценка автоматической семантической сегментации была выполнена на разработанном наборе данных SVAI. Предложенный алгоритм сравнивается

с современными автоматическими методами. Среднее пересечение-над-объединением (mIoU) использовалось в качестве меры точности семантической сегментации. Сравнение количественных значений mIoU для каждого из методов показано в табл. 3.

Метод	Набор данных SVAI
OneGAN	57,48
Voynov	69,12
Spectral Methods	75,81
TokenCut	72,36
Разработанный метод	79,91

Таблица 3 — Значения метрики mIoU на разработанном наборе данных для различных алгоритмов.

Качественные результаты семантической сегментации по предлагаемой нейросетевой модели показаны на рис. 9.

Предложенная архитектура демонстрирует эффективность при обработке аэрофотоснимков сложной структуры и обеспечивает высокую точность семантической сегментации за счет комбинации методов глубокого обучения и графовых подходов.

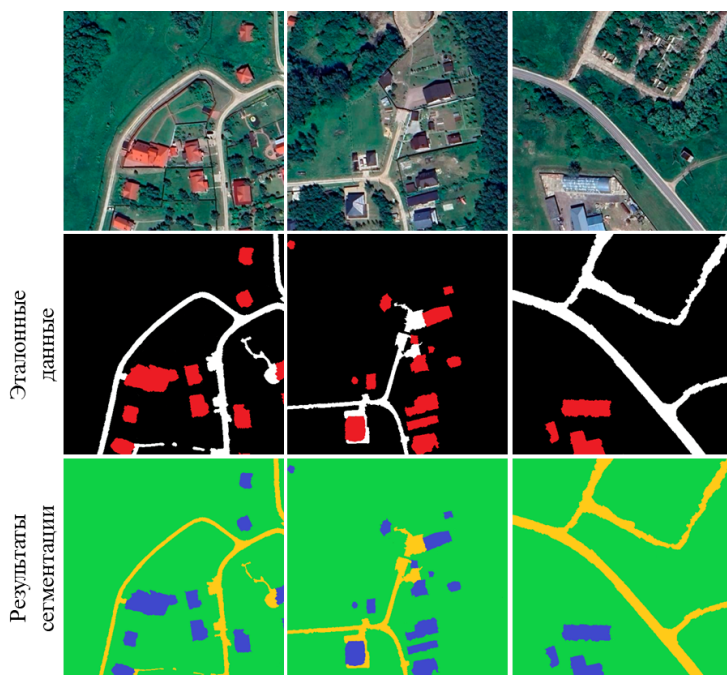


Рисунок 9 — Семантическая сегментация. Верхний ряд: исходное изображение; средний ряд: истинные данные; нижний ряд: результат работы предлагаемого алгоритма

В **заключении** приведены основные результаты работы, которые заключаются в следующем:

1. Разработан алгоритм на основе бинарной сегментации с использованием матрицы окрестности для упорядочивания прогнозируемых контуров зданий и улучшения их геометрии.
2. Разработан алгоритм автоматического извлечения границ зданий на основе бинарной сегментации с использованием многоэтапной регуляризации, включающей метод матрицы окрестности, оптимизацию спектральных границ и адаптивное уточнение контуров.
3. Разработан алгоритм векторизации аэрофотоснимков, основанный на обучении с подкреплением.

4. Разработана нейросетевая модель для точной сегментации аэрофотоснимков на основе визуального трансформера и графовой нейронной сети.
5. Создана представительная выборка, содержащая разновременные аэрофотоснимки различных участков земной поверхности.

Разработанные в исследовании методы сегментации и векторизации аэрофотоснимков вносят комплексный вклад в модернизацию процессов обновления геопространственной информации. Предложенные алгоритмы на основе бинарной сегментации с матрицей окрестности и многоэтапной регуляризации позволяют автоматизировать выявление пространственных изменений, включая детектирование новых объектов капитального строительства и трансформацию контуров существующих сооружений. Метод построения иерархического графа семантической разметки обеспечивает точное векторное представление данных с сохранением топологических отношений между объектами и учетом контекстных связей.

Существенное ускорение процессов обновления достигается за счет алгоритма на основе обучения с подкреплением, который имитирует логику оператора-дешифровщика и позволяет сократить временные затраты на обработку данных. Нейросетевая модель на основе визуального трансформера и графовой сети способствует улучшению детализации и структурирования данных через выделение объектов сложной геометрической формы и их классификацию по семантическим признакам. Применение единой методологической основы для сегментации и векторизации гарантирует согласованность пространственных данных и их соответствие стандартам ГИС.

Практическая значимость разработанных методов заключается в создании технологического фундамента для построения систем автоматизированного обновления геопространственной информации, способных оперативно отражать изменения на местности и поддерживать актуальность цифровых карт, кадастровых данных и тематических геоинформационных ресурсов.

Публикации автора по теме диссертации

1. *Emelyanov, A.* Extracting building outlines based on convolutional neural networks using the property of linear connectivity [Text] / A. Emelyanov, V. A. Knyaz, V. V. Kniaz // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. — 2024. — Vol. XLVIII-1—2024. — P. 147—152. — URL: <https://isprs-archives.copernicus.org/articles/XLVIII-1-2024/147/2024/>. — (Scopus, WoS, BAK).
2. Pixels relationship analysis for extracting building footprints [Text] / A. Emelyanov [et al.] // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. — 2024. — Vol. XLVIII-3—2024. — P. 141—146. — URL: <https://isprs-archives.copernicus.org/articles/XLVIII-3-2024/141/2024/>. — (Scopus, WoS, BAK).
3. Aerial Images Segmentation with Graph Neural Network [Text] / A. V. Emelyanov [et al.] // ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. — 2024. — Vol. X-2/W1—2024. — P. 1—8. — URL: <https://isprs-annals.copernicus.org/articles/X-2-W1-2024/1/2024/>. — (Scopus, WoS, BAK).
4. *Emelyanov, A.* Neural network algorithm based on reinforcement learning for vectorization of aerial and space images [Text] / A. Emelyanov, V. Knyaz, V. Kniaz // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. — 2025. — Vol. XLVIII-2/W9—2025. — P. 57—62. — URL: <https://isprs-archives.copernicus.org/articles/XLVIII-2-W9-2025/57/2025/>. — (Scopus, WoS, BAK).
5. Measuring Dental Enamel Thickness: Morphological and Functional Relevance of Topographic Mapping [Text] / A. V. Gaboutchian [et al.] // Journal of Imaging. — 2023. — Vol. 9, no. 7. — URL: <https://www.mdpi.com/2313-433X/9/7/127>. — (Scopus, WoS, BAK).

6. Hierarchical Scene Graph Generation and Vectorization of Aerial Images [Text] / V. A. Knyaz [et al.] // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. — 2025. — Vol. XLVIII-2/W9—2025. — P. 161—167. — URL: <https://isprs-archives.copernicus.org/articles/XLVIII-2-W9-2025/161/2025/>. — (Scopus, WoS, BAK).

Емельянов Антон Владимирович

Дешифрирование и векторизация аэро- и космофотоснимков методами машинного
обучения для обновления геопространственной информации

Автореф. дис. на соискание ученой степени канд. техн. наук

Подписано в печать ____ . ____ . ____ . Заказ № _____

Формат 60×90/16. Усл. печ. л. 1. Тираж 100 экз.

Типография _____