

УДК 004.932.75'1, 004.89

 10.25209/2079-3316-2025-16-2-111-152

Повышение точности сегментирования объектов с использованием генеративно-состязательной сети

Игорь Викторович **Винокуров** 

Финансовый Университет при Правительстве Российской Федерации, Москва, Россия

 igvvino@vinokurov@fa.ru

Аннотация. Маски, полученные с использованием модели глубокого обучения Mask R-CNN, в ряде случаев могут содержать фрагментированные контуры, неровные границы, ложные сращения соседних объектов и участки с пропущенной сегментацией. Чем больше объектов детектирования на изображении и меньше их размер, тем более чаще встречаются различного вида недостатки их масок. Примерами таких изображений могут являться аэрофотоснимки коттеджных и садовых товариществ и кооперативов, характеризующихся высокой плотности застройки. Для коррекции указанных недостатков предлагается использовать модель генеративно-состязательной сети, выполняющую постобработку предсказанных Mask R-CNN масок.

Качественная оценка сформированной в работе модели продемонстрировала, что она способна на приемлемом уровне восстанавливать целостность контуров, заполняет пропущенные области и разделять ошибочно объединенные объекты. Количественный анализ по метрикам IoU, precision, recall и F1-score показал статистически значимое улучшение качества сегментации по сравнению с исходными масками Mask R-CNN. Полученные результаты подтвердили, что предложенный подход позволяет довести точность формирования масок объектов до уровня, удовлетворяющего требованиям их практического применения в системах автоматизированного анализа аэрофотоснимков.

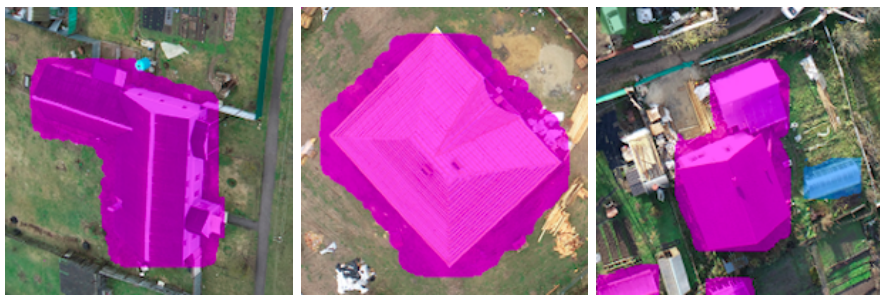
Ключевые слова и фразы: Компьютерное зрение, сегментация изображений, маски объектов, генеративно-состязательные сети, Mask R-CNN, PyTorch

Для цитирования: Винокуров И. В. *Повышение точности сегментирования объектов с использованием генеративно-состязательной сети* // Программные системы: теория и приложения. 2025. Т. 16. № 2(65). С. 111–152.
https://psta.psisras.ru/read/psta2025_2_111-152.pdf

Введение

Выявление точных масок объектов недвижимости на аэрофотоснимках с плотной застройкой представляет определённую сложность для моделей глубокого обучения, Mask R-CNN, YOLO и других. Основной причиной является близкое расположение зданий друг к другу, что затрудняет точное разделение объектов, поскольку их контуры могут пересекаться или сливаться на изображении.

Это особенно проблематично в условиях высокой детализации, где небольшие элементы архитектуры, такие как пристройки, балконы или общие стены между строениями, могут быть восприняты моделью как единый объект. Кроме того, сложная форма зданий, например, многоугольные очертания или нестандартные конструкции требует от модели более глубокого анализа и адаптации, что увеличивает вероятность ошибок при формировании масок. Несколько примеров таких масок приведены на рисунке 1.



(а) Неточная сегментация (б) Частичное отсутствие сегментации (в) Слияние масок объектов

Рисунок 1. Ошибки при формировании масок объектов

Ещё одной причиной является наличие внешних факторов, таких как растительность и тени, которые частично скрывают объекты на снимках. Деревья, кустарники или даже временные объекты, такие как автомобили, могут перекрывать части зданий, создавая разрывы в данных, доступных модели. Тени, вызванные различными углами съемки или временем суток, также искажают визуальную информацию, снижая контраст и затрудняя выделение четких границ объектов. Эти факторы приводят к тому, что модель может либо недооценивать площадь объекта, игнорируя скрытые участки, либо переоценивать ее, включая посторонние элементы в маску. В результате повышается потребность в дополнительной обработке и улучшении качества предсказаний, чтобы обеспечить корректность итоговых данных.

Проблема уточнения границ объектов особенно актуальна для информационной системы (ИС) ППК «Роскадастр», где к точности пространственных данных предъявляются повышенные требования. Традиционные методы постобработки, такие как морфологические операции или алгоритмы активных контуров, часто оказываются недостаточно эффективными для сложных случаев плотной застройки. В связи с этим возникает необходимость в решении задачи автоматического корректирования результатов сегментации, сохраняя при этом топологическую правильность объектов.

В представленной работе предлагается одно из возможных решений этой задачи, основанное на применении генеративно-сопоставительной сети (GAN) для постобработки результатов сегментации, полученных с помощью Mask R-CNN. Основная идея исследования заключается в том, что GAN-архитектура, должным образом обученная на репрезентативной выборке объектов застройки, характерной для ППК «Роскадастр», сможет эффективно восстанавливать целостность нарушенных границ, устранять внутренние разрывы масок, разделять ошибочно слившиеся объекты и сохранять точное положение угловых (поворотных) точек зданий.

Реализация исследований состояла из следующих этапов:

- (1) Формирование репрезентативного и сбалансированного датасета, содержащего соответствия исходных масок от Mask R-CNN эталонным маскам, извлечённым из файлов с аннотациями объектов и маскам, полученные в результате их экспертной коррекции.
- (2) Разработка архитектуры GAN, учитывающей особенности плотной застройки в коттеджных и садовых товариществах и кооперативах, требования к точности границ для кадастрового учета и необходимость сохранения топологии объектов.
- (3) Проведение экспериментальных исследований с оценкой качественных характеристик получаемых масок и количественных показателей (IoU, precision, recall, F1-score).

1. Обзор работ по использованию GAN для улучшения результатов распознавания объектов

В настоящее время существует значительное количество научных исследований, посвященных применению генеративно-сопоставительных сетей (GAN) для повышения эффективности распознавания изображений и их трансформации.

В работе [3] рассматриваются модели условно-сопоставительных сетей (conditional GAN, cGAN), которые могут быть применены для преобразования изображений одного типа в другой. Например, маски сегментации могут быть преобразованы в реалистичные изображения, или наоборот. Описанные модели демонстрируют высокую точность сегментации объектов и улучшают качество их масок.

Модель Pix2Pix, описанная в [4] и основанная на cGAN, используется для синтеза высококачественных изображений на основе входных данных, таких как маски сегментации. В статье показано, что GAN может улучшить качество масок путем обучения на парах «маска-точная маска».

В исследовании [5] представлена модель MaskGAN, предназначенная для интерактивного редактирования изображений лиц с использованием масок. Модель улучшает качество масок и позволяет манипулировать сегментированными областями изображений.

В [6] предлагается модель SegAN с многоуровневым дискриминатором и L1-потерями для медицинской сегментации. Модель использует многоуровневый дискриминатор для улучшения качества масок, особенно в задачах с размытыми границами объектов. GAN-модель улучшает границы объектов, снижает уровень шума и устраняет неточности детектирования объектов.

Механизм внимания для генерации масок в Attention-GAN, описанный в [7], применяется для трансформации объектов на фотографиях дикой природы. Модель позволяет устранять шумы и неточности детектирования, сохраняя мелкие детали объектов.

В статье [8] предлагается метод преобразования изображений с использованием CycleGAN и циклических потерь при обучении. Модель может быть адаптирована для улучшения масок сегментации, особенно в задачах с несоответствующими данными. Она устраняет шумы и неточности детектирования объектов, сохраняя их мелкие детали.

Модель Boundary-Aware GAN с многоуровневым дискриминатором, описанная в [9], предназначена для улучшения границ объектов в задачах семантической сегментации. Дискриминатор обучается различать реальные и предсказанные границы. Улучшения детектирования объектов заключаются в выявлении четких границ, снижении уровня шума и устранении неточностей.

В статье [10] исследуется применение GAN для бинарной семантической сегментации на несбалансированных датасетах. Авторы предлагают новую архитектуру и подход к обучению GAN для улучшения качества предсказания масок объектов в условиях сильной разбалансировки классов. Основное внимание уделяется проблемам низкой точности предсказания для миноритарного класса, потере мелких деталей объектов, шуму и артефактам в предсказанных масках. Авторы демонстрируют, что GAN позволяет улучшить качество масок за счёт более точного моделирования границ объектов и восстановления мелких деталей.

В работе [11] описывается модель PLGAN (Power-Line GAN), предназначенная для сегментации тонких и длинных объектов на фоне сложных текстур и шума. Основная цель — улучшить точность выделения линий электропередач, что важно для мониторинга инфраструктуры, картографирования и анализа окружающей среды. В статье предложены ключевые идеи, включая адаптацию GAN для сегментации тонких объектов, улучшение детализации, обработку несбалансированных данных и сглаживание контуров с использованием активных контуров и сплайн-интерполяции.

Проблема несбалансированного распределения данных между исходным доменом (совокупность изображений) и целевым доменом в задачах семантической сегментации аэрофотоснимков рассматривается в [12]. Авторы предлагают метод обучения модели без использования разметки для целевого домена. Для этого используется GAN, которая адаптирует стилистические особенности исходного домена к целевому. Основная цель этой работы — улучшить точность сегментации на целевом домене, минимизируя влияние различий между доменами, таких как изменения в освещении, ракурсы съемки или текстуры объектов. Авторы отмечают следующие улучшения, достигнутые с помощью предложенного подхода: лучшие обобщения модели, устранение влияния различий между доменами, сохранение мелких деталей и снижение шума.

В статье [13] исследуется применение cGAN для задач семантической сегментации спутниковых изображений среднего разрешения. Авторы фокусируются на улучшении точности выделения объектов на таких изображениях, которые часто страдают от шума, недостаточной детализации и сложных текстур. Основная цель работы — продемонстрировать, как cGAN может быть использована для повышения качества сегментации по сравнению с традиционными подходами, такими как U-Net или DeepLab. Особое внимание уделяется задачам мониторинга землепользования, городского планирования и анализа изменений в окружающей среде. Улучшениями, достигнутыми с помощью предложенного подхода, являются высокая точность для миноритарных классов, чёткие границы объектов, сохранение мелких деталей и устранение шума.

В [14] авторы предлагают новый подход, который использует GAN для улучшения качества предсказанных масок за счёт более точного моделирования границ объектов и восстановления мелких деталей. Основное внимание уделяется задачам, связанным с анализом изображений с высоким разрешением, таких как аэрофотоснимки, медицинские изображения и городские ландшафты. Цель работы — показать, что использование GAN может значительно повысить точность и детализацию

сегментации экземпляров объектов по сравнению с традиционными методами, такими как Mask R-CNN или DeepLab. Улучшение детализации и сглаживание контуров масок являются основными результатами этой работы.

2. Формирование набора данных

Использованный в экспериментальных исследованиях датасет представляет собой совокупность пар (предсказанная маска, точная маска). Предсказанная маска это бинарная маска объекта, полученная с использованием модели Mask R-CNN. Точная маска представляет собой бинарную маску, сформированную из полигона того же самого объекта, извлечённого из json-файла.

Типы и количества объектов, представленных в датасете, а также особенности формирования json-файлов к файлам с изображениями, описаны в [1]. Для отсутствующих в json-файлах полигонов объектов, соответствующие им точные маски формировались (корректировались) вручную. Это позволило создать репрезентативный и согласованный набор данных, включающий как автоматически сгенерированные, так и ручные маски. Все маски датасета были приведены к размеру 256×256 пикселей. Большой размер масок приводит к резкому увеличению вычислительной нагрузки и, соответственно, времени вычислений, меньший – к неровностям контуров предсказанных масок, после увеличения их размера. Увеличение до размера изображения с объектами недвижимости необходимо на этапе визуализации результата с целью совмещения маски объекта и содержащего его изображения.

3. Формирование и исследование модели

3.1. Особенности архитектуры GAN

GAN состоят из двух основных компонент: генератора и дискриминатора, которые работают в состязательном режиме. Основная цель GAN заключается в обучении генератора создавать данные, неотличимые от реальных.

Генератор в процессе обучения принимает случайный шум (обычно из нормального распределения) в качестве входных данных и преобразует его в некоторые данные, например, изображения. Задача генератора – сформировать данные, которые выглядят как реальные. Для этого он обучается минимизировать потери, возникающие при попытке дискриминатора определить подделку. Основными архитектурами генератора являются следующие:

DCGAN (Deep Convolutional GAN) [15]. Простой и эффективный, но может выдавать менее детализированные результаты на сложных данных. Генератор использует транспонированные свёртки для постепенного увеличения разрешения изображения. Начинает со случайного вектора (латентного пространства) и последовательно увеличивает размерность данных до целевого разрешения.

U-Net [16]. Подходит для задач, где важно сохранить пространственные детали, например, сегментация или улучшение масок. Генератор основан на архитектуре U-Net, которая включает энкодер для понижения размерности и декодер для восстановления изображения. Использует механизм пропусков (skip connections), что позволяет сохранять мелкие детали.

StyleGAN [17]. Высокая детализация и контроль над стилем, но требует больших вычислительных ресурсов. Позволяет контролировать различные уровни детализации изображения (например, общую форму и текстуры). Генератор разделен на две части для сопоставления стилей и для синтеза.

Дискриминатор оценивает качество сгенерированных данных. Задача дискриминатора — помочь генератору генерировать более правдоподобные данные (изображения), предоставляя корректные сигналы ошибки. Основные архитектуры дискриминатора:

PatchGAN [3]. Эффективен для задач, где важны границы объектов и мелкие детали. Анализирует изображение по небольшим локальным областям («патчам») вместо всего изображения. Каждый патч классифицируется как реальный или поддельный, что позволяет фокусироваться на локальных деталях.

Многоуровневый дискриминатор [18]. Улучшает качество генерации за счёт более комплексной оценки данных. Использует несколько дискриминаторов, каждый из которых работает на разных уровнях разрешения. Это позволяет оценивать как глобальные, так и локальные характеристики изображения.

Conditional GAN (cGAN) [19]. Наиболее эффективен для задач, где требуется направленная генерация, например, перевод изображений между их группами. Дискриминатор принимает дополнительную информацию (например, метки классов или условные данные) вместе с изображением, что позволяет контролировать тип генерируемых данных.

Учитывая особенности описанных выше архитектур генератора и дискриминатора, для повышения точности масок мелких объектов недвижимости были выбраны архитектуры U-Net и PatchGAN соответственно.

3.2. Формирование модели

Генератор, реализованный на основе архитектуры U-Net, реализует два основных этапа: сжатие (downsampling) и расширение (upsampling).

Этап сжатия уменьшает пространственное разрешение данных. Каждая из 5-ти групп элементов уменьшает размеры изображения в 2 раза и состоит из свёрточного слоя (`Conv2d`) с размером ядра (`kernel_size`) 4x4, шагом (`stride`) 2 и дополнением (`padding`) 1. За ним может следовать слой `BatchNorm2d` с функцией активации `LeakyReLU` и коэффициентом 0.2.

Элементы расширения наоборот, увеличивают пространственное разрешение данных. Элементов расширения в предлагаемом генераторе тоже 5. Каждый такой элемент увеличивает размеры изображения в 2 раза за счёт выполнения транспонированной свёртки (`ConvTranspose2d`) с размером ядра 4x4, шагом 2 и дополнением 1. Затем следует слой `BatchNorm2d` (кроме последнего блока) с функцией активации: `ReLU` (кроме последнего блока, где используется `Tanh` для ограничения выхода в диапазон $[-1, 1]$). Организация генератора GAN показана на рисунке 2.

Элемент расширения объединяет свои выходные данные с данными из соответствующего ему элемента сжатия изображения посредством скип-коннекторов. Это позволяет сохранять детали и улучшать качество генерации. На выходе генератора формируется одноканальное изображение размером, равным входному. Описанные выше элементы сжатия и расширения генератора 1-го уровня приведены на рисунках 3 и 4 соответственно.

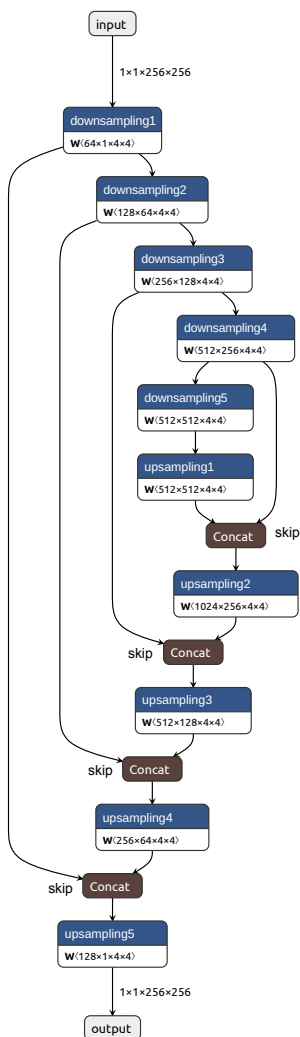


РИСУНОК 2. Генератор GAN

Как уже было отмечено выше, дискриминатор на основе архитектуры PatchGAN анализирует изображение по небольшим локальным областям (патчам). Каждый патч классифицируется им как реальный

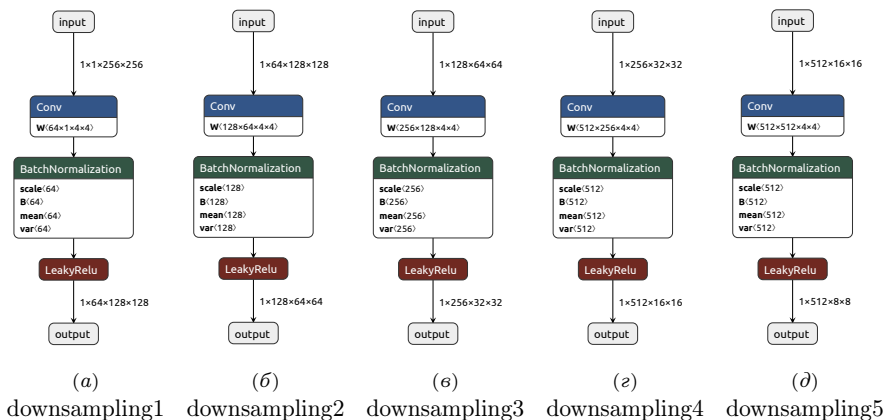


РИСУНОК 3. Элементы сжатия генератора GAN

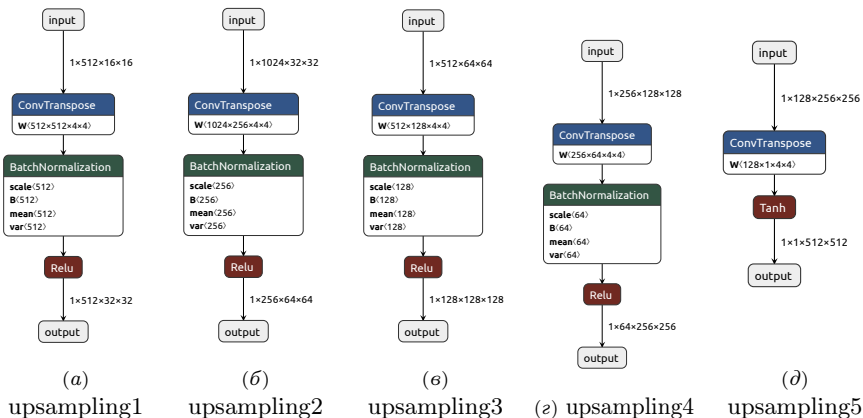


РИСУНОК 4. Элементы расширения генератора GAN

или поддельный, что позволяет фокусироваться на локальных деталях. Дискриминатор принимает два входных изображения: предсказанное изображение (сгенерированное генератором) и реальное изображение (в нашем случае они представляют собой точные маски объектов). Затем эти изображения объединяются по каналам. Последующие несколько свёрточных слоёв (Conv2d) с размером ядра 4×4 , шагом 2 и дополнением 1 уменьшают пространственные размеры данных.

Начиная со второго слоя применяется BatchNorm2d с функцией активации LeakyReLU и коэффициентом 0.2. Последний слой с функцией активации Sigmoid вычисляет вероятность того, что входные данные являются реальными. Структура дискриминатора приведена на рисунке 5.

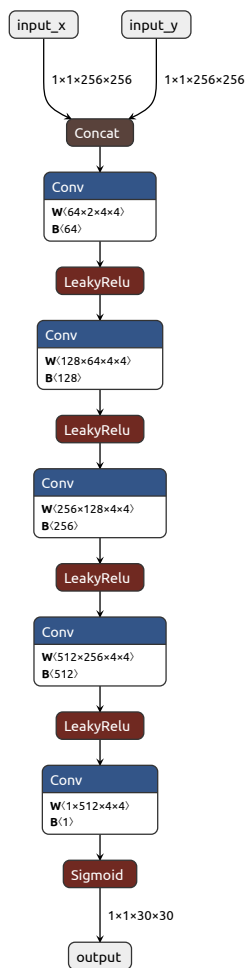


РИСУНОК 5. Дискриминатор GAN

В генераторе используется стандартная функция потерь для задач бинарной классификации `BCEWithLogitsLoss` и `L1 Loss` для минимизации пиксельной разницы между сгенерированными и реальными масками. Первая из этих функций используется для «обмана» дискриминатора, вторая помогает генератору создавать маски, которые близки к реальным по пиксельным значениям. Важным параметром при `L1 Loss` является λ_{L1} , который управляет важностью пиксельной разницы между сгенерированными и реальными масками. Это позволяет балансировать между «обман» дискриминатора и точностью воспроизведения деталей. Функция потерь генератора – `BCEWithLogitsLoss`.

Обучение GAN связано с определёнными проблемами, такими как нестабильность процесса обучения, коллапс мод и низкое качество генерации, что требует применения специальных методов для их решения. Одним из таких подходов является спектральная нормализация, которая ограничивает спектральную норму весов в слоях генератора и дискриминатора, тем самым стабилизируя обучение и предотвращая взрыв градиентов [20]. Другой метод заключается в использовании штрафа за градиент (gradient penalty), который добавляет регуляризацию к выходам дискриминатора, обеспечивая более плавное распределение значений и улучшая сходимость модели [21]. Также широкое применение нашёл механизм самовнимания (self-attention), позволяющий модели учитывать глобальные зависимости в данных, что особенно полезно при работе с высокоразмерными изображениями [22]. Однако, поскольку восстановление простых бинарных масок является относительно простой задачей, и как показали результаты исследования описанной выше модели GAN, использование этих подходов в данном случае является излишним.

3.3. Результаты исследования модели

Формирование и исследование модели GAN проводилось с использованием фреймворка `PyTorch` в среде `Google Colab Pro`. `Jupyter`-ноутбук (лицензия MIT) [доступен по этой ссылке](#):  `jupyter`. Исследования основных метрик точности были выполнены для различных значений скорости обучения, количества эпох и размеров батча. Метрики точности генератора

GAN исследовались при $\lambda_{L1} = 100$, поскольку при таком значении их кривые обладают достаточной степенью гладкости, что свидетельствует об отсутствии взрывных или затухающих градиентов. На рисунках 6, 7 и 8 показаны кривые метрик точности для значений $\lambda_{L1} = 80, 100$ и 120 .

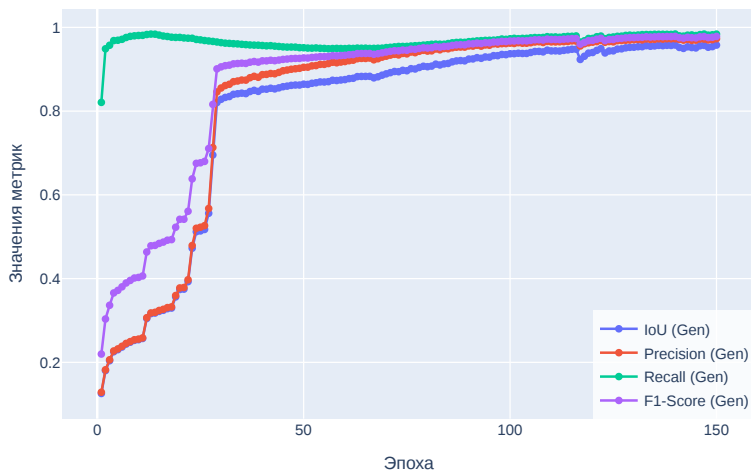


Рисунок 6. Графики метрик точности генератора GAN для $\lambda_{L1}=80$

В таблице 1 приведены результаты исследований основных метрик точности IoU, IoU0.5-95, Precision, Recall и F1-Score. Как показали проведённые исследования, использование GAN позволяет значительно улучшить точность масок сегментации объектов недвижимости, полученных с использованием модели Mask R-CNN [1].

Рассмотрим несколько примеров улучшения масок объектов недвижимости, полученных с использованием модели Mask R-CNN. На рисунке 9 показаны маски, предсказанные моделью Mask R-CNN, и улучшенные маски, сгенерированные с помощью GAN. По значениям метрик точности можно сделать вывод, что использование GAN позволяет значительно улучшить качество масок за счёт коррекции ошибок, таких как разрывы контуров, шум и неточности в границах объектов.

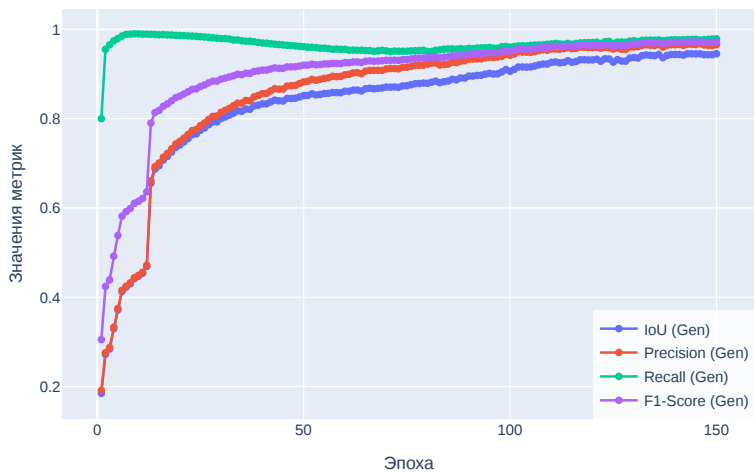


Рисунок 7. Графики метрик точности генератора GAN для $\lambda_{L1}=100$

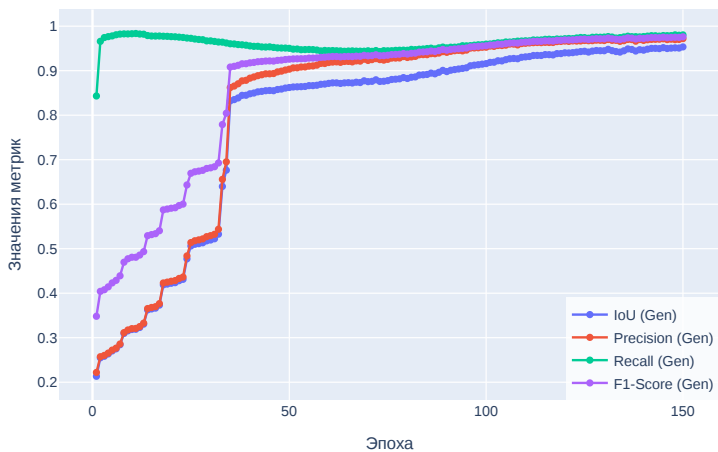


Рисунок 8. Графики метрик точности генератора GAN для $\lambda_{L1}=120$

Таблица 1. Оценка метрик точности для предсказанных GAN масок в зависимости от параметров обучения модели

Скорость обучения	Размер батча	Эпох обучения	IoU (Pred)	Precision (Pred)	Recall (Pred)	F1-Score (Pred)	IoU0.5-95 (Pred) (Mask R-CNN)	IoU0.5-95 (Pred) (GAN)
1e-6	10	500	0.87	0.92	0.91	0.91	0.67	0.90
5e-6	10	400	0.85	0.90	0.89	0.90	0.65	0.88
5e-6	20	350	0.84	0.89	0.88	0.88	0.64	0.87
1e-5	10	300	0.83	0.88	0.87	0.87	0.63	0.86
1e-5	20	250	0.84	0.89	0.88	0.88	0.64	0.87
5e-5	8	200	0.79	0.84	0.83	0.83	0.59	0.81
5e-5	10	300	0.82	0.87	0.86	0.86	0.62	0.85
5e-5	16	250	0.81	0.86	0.85	0.85	0.61	0.84
4e-5	8	150	0.80	0.85	0.84	0.84	0.60	0.82
4e-5	16	200	0.78	0.83	0.82	0.82	0.58	0.80
1e-4	8	100	0.75	0.80	0.78	0.79	0.55	0.78
1e-4	16	150	0.77	0.82	0.81	0.81	0.57	0.80
1e-4	20	200	0.78	0.83	0.82	0.82	0.58	0.81
2e-4	8	80	0.70	0.75	0.73	0.74	0.50	0.73
2e-4	16	100	0.72	0.78	0.76	0.77	0.52	0.76



Рисунок 9. Результаты улучшения с помощью GAN показаны под предсказанными Mask R-CNN масками объектов недвижимости

Пример улучшения масок объектов на аэрофотоснимке, на которых сложность задачи возрастает из-за высокой плотности объектов и их разнообразия приведён на рисунках 10.

При генерации масок с помощью GAN часто возникает проблема ступенчатых краёв, см. рисунок 9. Это обусловлено спецификой работы GAN, которая связана с дискретностью пикселей при формировании изображений и сложностью точного воспроизведения мелких деталей границ объектов. Для улучшения контуров масок сегментации могут быть применены различные функции морфологического сглаживания, которые позволяют устранить шум, заполнить пробелы и сделать границы объектов более гладкими.

Одним из базовых подходов является использование операций морфологического открытия и закрытия с заданным структурирующим элементом (например, квадратной матрицей (5×5)). Открытие помогает удалить мелкие шумовые артефакты, а закрытие заполняет мелкие разрывы внутри маски. Для дополнительного сглаживания контуров



(a) Обнаруженные моделью Mask R-CNN маски объектов недвижимости



(б) Маски улучшены генератором GAN

РИСУНОК 10. Улучшение масок объектов недвижимости

можно применить Гауссово размытие (**GaussianBlur**), которое смягчает резкие переходы между фоном и объектом.

Также эффективным методом является использование активных контуров (**snakes**), которые адаптивно подстраиваются под форму объекта за счёт минимизации энергии упругости и гладкости. Для достижения еще более гладких контуров может быть применена сплайн-интерполяция, которая аппроксимирует исходный контур с помощью кубических сплайнов, обеспечивая равномерное распределение точек и плавность границ.

Ещё одним из эффективных способов устранения ступенчатости является использование алгоритма Дугласа-Пекера, который специализируется на упрощении полилиний и сглаживании контуров [23]. Особенность данного алгоритма заключается в его способности анализировать каждую точку границы маски и оценивать ее значимость для общей формы объекта. Точки, которые вносят минимальный вклад в конфигурацию контура, удаляются, что позволяет получить более плавные и естественные границы без потери ключевых геометрических особенностей. Пример использования алгоритма Дугласа-Пекера приведён на рисунке 11.

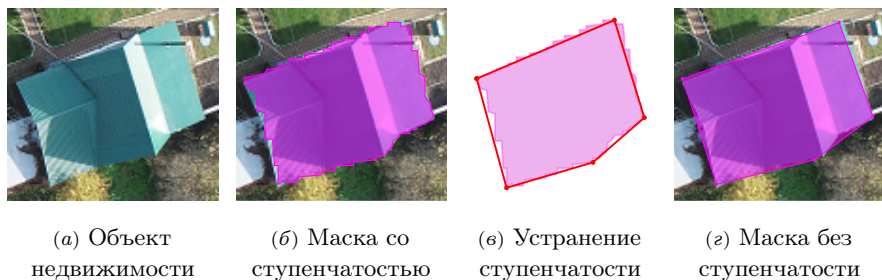


РИСУНОК 11. Устранения ступенчатости маски с использованием алгоритма Дугласа-Пекера

Эти методы или их комбинация позволяют в ряде случаев улучшить качество масок, делая их более пригодными для последующего анализа или визуализации. Как показали проведённые исследования, использования методов морфологического сглаживания и улучшения границ масок с использованием алгоритма Дугласа-Пекера в постобработке масок, даёт возможность повысить точность их формирования в среднем на 5–7%, что особенно важно для задач компьютерного зрения и сегментации объектов на аэрофотоснимках в частности.

Заключение

В данной работе была предложена и исследована модель GAN, предназначенная для повышения точности масок сегментации, полученных с использованием модели Mask R-CNN. Эксперименты показали, что применение GAN позволяет существенно улучшить качество детектированных масок объектов недвижимости за счёт коррекции ошибок, таких как разрывы в масках, шумовые артефакты и неровные контуры. Улучшение достигается за счёт использования функции потерь L1 для минимизации пиксельных различий между целевыми и сгенерированными масками, а также за счёт работы дискриминатора, который способствует более реалистичной генерации масок.

Результаты проведенных исследований демонстрируют улучшение метрик точности, таких как IoU, Precision, Recall и F1-Score, для улучшенных масок по сравнению с исходными предсказаниями Mask R-CNN. Среднее значение IoU_{0.5-95} увеличилось в среднем на 30-35%, что подтверждает эффективность предложенного подхода. Дополнительно было отмечено, что использование морфологических операций и методов сглаживания контуров позволит дополнительно улучшить гладкость и точность границ масок.

Полученные результаты имеют важное практическое значение и будут использованы в ИС ППК «Роскадастр» для автоматического преобразования растровых изображений с объектами недвижимости в векторные карты. Это позволит повысить точность и скорость формирования цифровых карт, что особенно важно для задач кадастрового учета и управления территориями. Предложенный подход может быть адаптирован для других приложений, связанных с обработкой изображений и компьютерным зрением, где требуется высокая точность сегментации объектов. Проведённое в работе исследование подтверждает перспективность применения GAN с целью улучшения качества масок объектов недвижимости, что позволяет его эффективно использовать для автоматизации процессов анализа и преобразования данных в геоинформационных системах. Дальнейшие исследования могут быть направлены на оптимизацию параметров модели, внедрение новых методов постобработки масок и расширение области применения GAN.

Список использованных источников

- [1] Винокуров И. В. *Использование модели Mask R-CNN для сегментации объектов недвижимости на аэрофотоснимках* // Программные системы: теория и приложения.– 2025.– Т. 16.– № 1(64).– С. 3–44.  [URL](#)  [↑136, 143](#)
- [2] G. Cohen, R. Giryes *Generative adversarial networks*.– 2024.– 28 pp.  [arXiv](#)  [2203.00667](#)  ↑
- [3] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros *Image-to-image translation with conditional adversarial networks*.– 2016.– 17 pp.  [arXiv](#)  [1611.07004](#)  [↑133, 137](#)
- [4] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, B. Catanzaro *High-resolution image synthesis and semantic manipulation with conditional GANs*.– 2017.– 14 pp.  [arXiv](#)  [1711.11585](#)  [↑134](#)
- [5] C.-H. Lee, Z. Liu, L. Wu, P. Luo *MaskGAN: Towards diverse and interactive facial image manipulation*.– 2019.– 20 pp.  [arXiv](#)  [1907.11922](#)  [↑134](#)
- [6] Y. Xue, T. Xu, H. Zhang, L. Rodney Long, X. Huang *SegAN: Adversarial network with multi-scale L_1 loss for medical image segmentation*.– 2017.– 9 pp.  [arXiv](#)  [1706.01805](#)  [↑134](#)
- [7] X. Chen, C. Xu, X. Yang, D. Tao *Attention-GAN for object transfiguration in wild images*.– 2018.– 18 pp.  [arXiv](#)  [1803.06798](#)  [↑134](#)
- [8] J.-Y. Zhu, T. Park, P. Isola, A. A. Efros *Unpaired image-to-image translation using cycle-consistent adversarial networks*.– 2017.– 18 pp.  [arXiv](#)  [1703.10593](#)  [↑134](#)
- [9] J. Gong, J. Xu, X. Tan, J. Zhou, Y. Qu, Y. Xie, L. Ma *Boundary-aware geometric encoding for semantic segmentation of point clouds*.– 2021.– 9 pp.  [arXiv](#)  [2101.02381](#)  [↑134](#)
- [10] L. Xu, M. Gabbouj *Revisiting generative adversarial networks for binary semantic segmentation on imbalanced datasets*.– 2024.– 14 pp.  [arXiv](#)  [2402.02245](#)  [↑134](#)
- [11] R. Abdelfattah, X. Wang, S. Wang *JPLGAN: Generative adversarial networks for power-line segmentation in aerial images*.– 2022.– 11 pp.  [arXiv](#)  [2204.07243](#)  [↑135](#)
- [12] B. Benjdira, Y. Bazi, A. Koubaa, K. Ouni *Unsupervised domain adaptation using generative adversarial networks for semantic segmentation of aerial images* // Remote Sens..– 2019.– Vol. 11.– No. 11.– id. 1369.– 23 pp.  [arXiv](#)  [1905.03198](#)  [↑135](#)
- [13] A. Kulkarni, T. Mohandoss, D. Northrup, E. Mwebaze, H. Alemohammad *Semantic segmentation of medium-resolution satellite imagery using conditional generative adversarial networks*.– 2020.– 7 pp.  [arXiv](#)  [2012.03093](#)  [↑135](#)
- [14] Q. H. Le, K. Youcef-Toumi, D. Tsetserukou, A. Jahanian *GAN Mask R-CNN: Instance semantic segmentation benefits from generative adversarial networks*.– 2020.– 13 pp.  [arXiv](#)  [2010.13757](#)  [↑135](#)
- [15] A. Radford, L. Metz, S. Chintala *Unsupervised representation learning with deep convolutional generative adversarial networks*.– 2016.– 16 pp.  [arXiv](#)  [1511.06434](#)  [↑137](#)
- [16] O. Ronneberger, Ph. Fischer, Th. Brox *U-Net: Convolutional networks for biomedical image segmentation*.– 2015.– 8 pp.  [arXiv](#)  [1505.04597](#)  [↑137](#)

- [17] T. Karras, S. Laine, T. Aila *A style-based generator architecture for generative adversarial networks.*– 2018.– 12 pp.  arXiv  1812.04948 ↑¹³⁷
- [18] T. Karras, T. Aila, S. Laine, J. Lehtinen *Progressive growing of GANs for improved quality, stability, and variation.*– 2017.– 26 pp.  arXiv  1710.10196 ↑¹³⁷
- [19] M. Mirza, S. Osindero *Conditional generative adversarial nets.*– 2014.– 7 pp.  arXiv  1411.1784 ↑¹³⁷
- [20] T. Miyato, T. Kataoka, M. Koyama, Y. Yoshida *Spectral normalization for generative adversarial networks.*– 2018.– 26 pp.  arXiv  1802.05957 ↑¹⁴²
- [21] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. C. Courville *Improved training of Wasserstein GANs.*– 2017.– 20 pp.  arXiv  1704.00028 ↑¹⁴²
- [22] H. Zhang, I. Goodfellow, D. Metaxas, A. Odena *Self-attention generative adversarial networks.*– 2019.– 10 pp.  arXiv  1805.08318 ↑¹⁴²
- [23] H. Chen *An improved Douglas-Peucker algorithm applied in coastline generalization*, Fourth International Conference on Geology, Mapping, and Remote Sensing (ICGMRS 2023) (14–16 April 2023, Wuhan, China), Proc. SPIE.– vol. **12978**.– 2024.– id. 1297804.  ↑¹⁴⁸

Поступила в редакцию 21.04.2025;
одобрена после рецензирования 10.06.2025;
принята к публикации 11.06.2025;
опубликована онлайн 28.06.2025.

Рекомендовал к публикации

д.ф.-м.н. А. М. Елизаров

Информация об авторе:



Игорь Викторович Винокуров

Кандидат технических наук (PhD), ассоциированный профессор в Финансовом Университете при Правительстве Российской Федерации. Область научных интересов: информационные системы, информационные технологии, технологии обработки данных



0000-0001-8697-1032

e-mail: igvvinokurov@fa.ru

Декларация об отсутствии личной заинтересованности: *благополучие автора не зависит от результатов исследования.*