

Применение GAN для улучшения разрешения изображений

Шерыхалина Н.М.
профессор, д.т.н.

Институт информатики, математики и
робототехники
Уфимский университет науки и технологий
Уфа, Россия Город
e-mail: n_sher@mail.ru

Савин Е.В.
студент

Институт информатики, математики и
робототехники
Уфимский университет науки и технологий
Уфа, Россия Город
e-mail: n778hr@mail.ru

Аннотация¹

В последние годы технологии генеративно-состязательных сетей (GAN) активно применяются для улучшения разрешения изображений. Они позволяют не только увеличивать разрешение, но и восстанавливать утраченные детали, что имеет огромное значение в таких областях, как медицина, видеонаблюдение и спутниковая съемка.

В данной работе проведено исследование принципов работы GAN в задачах суперразрешения изображений. В него включено рассмотрение наиболее популярных на сегодняшний день моделей, таких как SRGAN и ESRGAN, а также обсуждение их текущих и перспективных применений.

В статье проводится обзор существующих архитектур GAN, их применения для суперразрешения и улучшения качества изображений. Рассматриваются ключевые улучшения в модели ESRGAN, включая использование перцептуальных потерь. Также приводятся примеры успешного применения данных технологий.

Проведенные исследования показывают, что использование GAN, особенно моделей SRGAN и ESRGAN, позволяет существенно улучшить визуальное качество изображений, особенно в медицинских и видеонаблюдательных системах. Модели продемонстрировали высокую эффективность в восстановлении текстур и мелких деталей.

Ключевые слова: суперразрешение изображений; генеративно-состязательные сети; SRGAN; ESRGAN; улучшение качества изображений

1. Введение

Качество изображения играет ключевую роль в таких областях, как медицина, промышленный дизайн. Для повышения качества изображения применяются методы интерполяции и экстраполяции — техники обработки данных, которые помогают восполнить пробелы в информации и улучшить визуальные характеристики изображения [1]. Однако на практике возникает проблема низкого разрешения изображений, что снижает их точность и уменьшает информационную плотность. На помощь приходят методы суперразрешения, которые позволяют повысить разрешение и восстановить детали изображения. Можно отметить значительный скачок в прогрессе в 2017 году, благодаря появлению модели SRGAN (Super-Resolution GAN). Эта модель стала первой, которая смогла эффективно увеличить разрешение изображений, сохраняя при этом видимую четкость тонких деталей и текстур [2].

В 2014 году произошел знаковый момент в развитии данной технологии, введение генеративных состязательных сетей (GAN). Эта концепция, предложенная Иэном Гудфеллоу и его коллегами, была далеко не обычным предложением. Состоящие из двух взаимозависимых нейронных сетей — генератора и дискриминатора — GAN сыграли важную роль в улучшении качества создаваемых изображений, хотя сама формулировка может показаться несколько громоздкой, учитывая техническую природу вопроса [3]. Действительно, можно утверждать, что такие достижения заложили основу для современных методов суперразрешения изображений.

Развитие технологии не остановилось на SRGAN. В 2018 году была представлена ESRGAN — усовершенствованная модель Super-Resolution GAN. Качество изображения значительно улучшилось, как и его реализм по сравнению с предыдущими версиями.

Труды X Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 12-14 ноября, Уфа-Баку-Чандигарх, 2024

Эти достижения в области усовершенствованного суперразрешения открыли новые возможности для его применения. Области, такие как медицина, спутниковая съемка и видеонаблюдение, являются одними из тех, где высокое качество изображения имеет критическое значение для точности анализа и диагностики, и именно в этих сферах достижения суперразрешения принесли значительные преимущества.

Принципы работы GAN и суперразрешения изображений

Что такое суперразрешение изображений

Суперразрешение — это технология, которая улучшает качество изображений, преобразуя исходные изображения с низким разрешением в более детализированные и качественные. Изображения с низким разрешением, часто представленные матрицами размером $n \times n$, преобразуются в версии с более высоким качеством, где число пикселей увеличивается до $k \times k$, где $k > n$, что позволяет восстанавливать утраченные детали. Это достигается благодаря добавлению более детализированной информации, основанной на сложных математических моделях, таких как интерполяция, свёрточные нейронные сети, что и делает возможным повышение качества.

Суперразрешение находит применение в самых разных областях: медицине, спутниковой съемке, обработке фото и видео, а также в системах видеонаблюдения, где особенно важна высокая точность [6]. С его помощью восстанавливаются детали, утраченные при снижении разрешения, и это становится особенно важным в ситуациях, когда невозможно получить исходные данные с высоким разрешением, например, при съемке на больших расстояниях или в условиях плохого освещения.

На сегодняшний день суперразрешение стало сложной и многогранной областью. Машинное обучение привело к созданию множества новых методов, которые оказались весьма эффективными. Например, генеративные состязательные сети (GAN) значительно улучшили качество восстанавливаемых изображений по сравнению с традиционными методами [5]. Тем не менее, классические методы интерполяции, такие как билинейная и бикубическая интерполяция [7], по-прежнему увеличивают количество пикселей, но они не способны восстановить тонкие и сложные детали. Это часто приводит к размытию изображений и потере важных элементов. В ситуациях, когда требуется значительное повышение разрешения, такие методы оказываются недостаточно эффективными и не могут полностью удовлетворить требования к качеству [5].

Анализ эффективности GAN

Генеративно-состязательные сети (GANs) зарекомендовали себя как эффективный инструмент для решения задач суперразрешения, позволяя не

только восстанавливать утраченные пиксели, но и воссоздавать текстуры, которые практически не отличаются от оригинальных. В этих сетях выделяются два основных компонента: генератор и дискриминатор. Генератор создает изображение с высоким разрешением, а дискриминатор оценивает, насколько оно похоже на реальное [5].

Без сомнения, работа GAN требует значительных вычислительных ресурсов. Обработка данных с высоким разрешением делает процесс обучения крайне ресурсоёмким, так как необходимо проанализировать большое количество изображений на каждой итерации. В последние годы исследования сосредоточены на снижении вычислительных затрат. Некоторые оптимизируют архитектуры моделей, например, с помощью легковесных сверточных сетей (lightweight CNNs), которые уменьшают количество операций свёртки, снижая тем самым вычислительную сложность с $O(n^2)$ до $O(n)$ для некоторых задач, а другие используют методы сжатия моделей [18]. Методы, включающие сокращение параметров и квантование, значительно снижают вычислительную нагрузку на системы, при этом качество изображения сохраняется [19]. Область квантования — это разрядность битов, или, точнее, её уменьшение. Снижение разрядности с 32 бит до всего лишь 8 бит упрощает вычисления. Вычислительная нагрузка существенно сокращается. Также есть методы сокращения параметров, такие как обрезка нейронных связей (pruning), при которых наименее значимые веса обнуляются. Это уменьшает количество необходимых вычислений, снижая сложность операций с $O(n^2)$ до $O(n \log n)$, но при этом качество изображения остаётся на достойном уровне.

Зависимость от обучающих данных

Работа GANs во многом определяется качеством и разнообразием обучающих данных. Если предоставленные модели данные недостаточно разнообразны или содержат артефакты, сгенерированные изображения могут оказаться искаженными или лишенными текстур. Это особенно критично для медицинских изображений или спутниковой съемки, где каждая деталь может быть иметь существенное значение для точной диагностики или анализа [14].

Чтобы справиться с этой проблемой, разработаны подходы, такие как аугментация данных, которые искусственно увеличивают объем обучающего набора, а также методы самообучения, которые позволяют моделям учиться на неразмеченных данных. Например, современные модели, такие как SRFlow, используют самообучение для улучшения качества суперразрешения изображений, уменьшая

при этом зависимость от заранее размеченных данных [21].

Проблема артефактов

Несмотря на значительные достижения в области суперразрешения, одной из главных сложностей использования GAN остается появление артефактов на изображениях. Визуальные искажения могут возникать, когда модель не может корректно восстановить мелкие детали или текстуры в высокоразрешённых изображениях. Для решения этих вопросов современные исследования предлагают использовать улучшенные функции потерь, такие как perceptual loss, которые фокусируются на визуальных характеристиках изображения, воспринимаемых человеком [22]. Эти подходы позволяют уменьшить размытость и искажения, делая изображения более естественными.

Будущие исследования

Текущие направления исследований сосредоточены на улучшении архитектуры GAN для суперразрешения, снижении вычислительных затрат и устранении артефактов. Например, внедрение более сложных функций потерь и оптимизация архитектур свёрточных сетей продолжают активно исследоваться [20]. Использование GAN для улучшения качества видео в реальном времени, в том числе в таких приложениях, как видеоконференции и онлайн-стриминг, представляет собой новое направление, которое также требует глубокого анализа и исследований [19].

Роль GAN в улучшении разрешения изображений

Генеративно-сопоставительные сети (GAN) стали революцией в задаче суперразрешения изображений благодаря их способности восстанавливать сложные детали и текстуры. Архитектура GAN включает две сети: генератор и дискриминатор. Генератор создает изображение с более высоким разрешением на основе изображения с низким разрешением, в то время как дискриминатор пытается отличить искусственно созданное изображение от реального [5]. Этот процесс позволяет генератору «учиться» создавать изображения, которые всё более похожи на реальные, постепенно улучшая качество синтезируемых данных.

Этот процесс можно представить как минимакс-игру между генератором G и дискриминатором D , где задача генератора — минимизировать вероятность того, что дискриминатор обнаружит синтетическое изображение, а задача дискриминатора — максимизировать свою способность правильно отличить синтетическое изображение от реального:

$$\min_G \max_D V(D, G) = E_{x \sim P_{data}(x)} [\log D(x)] + E_{z \sim P_z} [\log(1 - D(G(z)))]$$

Здесь

- x — реальное изображение, взятое из распределения P_{data}
- z — случайный шум, взятый из распределения P_z ,
- $G(z)$ — сгенерированное изображение.

Этот подход был предложен в оригинальной работе Иэна Гудфеллоу и его команды в 2014 году [5]. Современные исследования продолжают развивать эту архитектуру, добавляя новые механизмы для повышения стабильности и качества генерации изображений [10].

Кроме того, современные исследования в области GAN сосредоточены на оптимизации архитектуры и снижении вычислительных затрат. Например, методы сжатия моделей, такие как квантование, уменьшают количество вычислений, что позволяет эффективно использовать модели в реальном времени, что особенно важно для таких приложений, как видеонаблюдение и медицинская диагностика. Эти техники сжатия применяются без значительных потерь в качестве изображения, что делает их перспективными для будущих разработок.

Структура GAN для задачи суперразрешения

Для задачи суперразрешения GAN используется в таких моделях, как SRGAN и ESRGAN, где генератор восстанавливает изображения высокого разрешения из входных данных низкого качества. Одной из ключевых характеристик этих моделей является использование функции перцептуальных потерь (perceptual loss), которая добавляет в процесс обучения высокоуровневые визуальные характеристики. В отличие от простого сравнения пикселей, перцептуальная потеря позволяет модели учитывать, как воспринимаются изображения нейронной сетью.

Формула перцептуальных потерь выглядит следующим образом:

$$L_{perceptual}(I_{HR}, I_{SR}) = \sum_{i=1}^n \|\phi_i(I_{HR}) - \phi_i(I_{SR})\|_2^2$$

Здесь

- I_{HR} — изображение с высоким разрешением (high resolution),
- I_{SR} — сгенерированное изображение с повышенным разрешением (super-resolution),
- ϕ_i — активации i -го слоя предобученной нейронной сети (например, VGG).

Использование перцептуальной потери позволяет GAN ориентироваться не только на пиксельные отличия между реальными и сгенерированными изображениями, но и на то, как эти изображения воспринимаются нейронной сетью, что делает их более реалистичными [4]. Более того, такие усовершенствования в функции потерь позволяют

GAN достигают более высокого уровня генерации деталей, недоступного для классических методов.

Пример успешного применения GAN для суперразрешения

Ярким примером успешного применения GAN для повышения разрешения изображений стал проект, реализованный в 2020 году. Целью этого проекта было улучшение качества медицинских снимков, включая МРТ и КТ. Исследования показали, что с помощью GAN можно значительно улучшить детализацию и точность интерпретации изображений. Это дало медицинским специалистам возможность более точно интерпретировать диагностические результаты. Визуализация мелких патологий и редких опухолей получила особое внимание благодаря этому подходу. Ключевая роль здесь заключается в содействии раннему выявлению заболеваний, что делает эту технологию незаменимой.

GAN также активно используется в системах видеонаблюдения. В эксперименте, проведенном в Москве, алгоритмы суперразрешения применялись для улучшения качества изображений с камер наблюдения. Это позволило более точно распознавать лица на удаленных объектах и повысить эффективность работы системы безопасности [10]. В Китае такие технологии активно развиваются в рамках проектов «Умный город». С их помощью удаётся значительно повысить уровень общественной безопасности в местах с большим скоплением людей [12].

GAN широко применяется и в восстановлении старых видеоматериалов и фотографий. Например, в проектах по реставрации старых фильмов GAN используется для увеличения разрешения, а также для устранения шумов и артефактов, что позволяет создавать высококачественные изображения для цифрового архивирования и повторного выпуска. Эти усилия особенно важны для сохранения цифрового наследия. Исследования в этой области продолжают открывать новые возможности, что делает восстановление старых видеоматериалов и фотографий удивительно обширной темой в мире современных технологий.

Исследования по применению генеративных состязательных сетей (GANs) для повышения качества изображений продолжаются, появляются перспективы для различных научных и коммерческих инициатив. Менее очевидным примером является использование современных моделей, известных как GAN-TTS (сокращение от Generative Adversarial Networks для суперразрешения временных рядов), которые применяются для анализа временных рядов и увеличения их разрешения. Данный подход может

иметь значение, особенно при глубокой обработке финансовых и метеорологических данных [13].

2. Примеры использования GAN для улучшения разрешения изображений

Распространение технологии GAN в последние годы заметно во множестве областей. Одним из ярких примеров является использование ESRGAN для повышения качества спутниковых и аэрофотоснимков. Исследования указывают, что GAN не только увеличивают разрешение изображений, но и помогают восстановить утраченные текстуры, что имеет большое значение при изучении сложных структур земной поверхности. Улучшение качества изображений в системах мониторинга окружающей среды и геодезии способствует повышению точности прогнозов природных катастроф [11].

Использование генеративно-состязательных сетей в медицинской визуализации также весьма важно и порой даже необходимо. SRGAN и ESRGAN, данные сети демонстрируют высокую эффективность в улучшении качества МРТ и КТ изображений. Это улучшение помогает врачам точнее интерпретировать диагностические результаты, что в конечном итоге приводит к повышению качества диагностики и лечения пациентов. Исследование 2021 года показало значительное увеличение точности диагностики при использовании метода суперразрешения на основе GAN, в сочетании с процессом глубокого обучения [12].

Старые фильмы и видеоматериалы обретают вторую жизнь благодаря технологиям GAN. SRGAN помогает улучшить качество старых фильмов, преобразуя их в такие форматы, как UHD и даже 4K. Алгоритмы не только увеличивают разрешение, но и восстанавливают утраченные текстуры, делая старые фильмы более интересным для новой и молодой аудитории [8].

Современные исследования показывают, что GAN активно применяются для улучшения качества изображений в реальном времени. Примером может служить сфера видеонаблюдения. С помощью инструментов, таких как ESRGAN, можно улучшить разрешение старых записей; основное применение технологии — распознавание лиц и объектов в условиях сложного освещения или низкого качества видео. В 2022 году провели исследование и оно показало, что генеративно-состязательные сети повысили точность распознавания образов на видео до 95% - это значительно превзошло все предыдущие методы [10].

Технология Generative Adversarial Network продемонстрировала сильный потенциал в улучшении качества изображений в реальном времени. Особенно роль заметна в таких областях, как беспилотные летательные аппараты и автономные транспортные средства. Генеративно-состязательные сети так же могут помочь в улучшения качества изображений с

дронов, тем самым помогая им распознавать лучше объекты по более качественному изображению [11].

3. Перспективы и будущее GAN в улучшении разрешения изображений

Текущие направления исследований

В последнее время исследования в области суперразрешения сосредоточили свое внимание на улучшении архитектуры моделей и снижении требований для вычислительных ресурсов. Например, такая сложная модель как Progressive Generative Adversarial Network (ProGAN) продемонстрировала возможность генерации изображений высокого разрешения при более стабильном процессе обучения [5]. Ученые исследуют новые способы по оптимизации архитектуры и так же разрабатывают более эффективные методы обучения для ускорения процесса получения изображения.

Одно из актуальных направлений сейчас – это разработка моделей способных обрабатывать видео в реальном времени. Однако обработка видео является более сложной задачей с точки зрения вычислительных мощностей, так как надо учитывать не только каждый кадр, но и связь между ними. Это увеличивает вычислительную сложность задачи в n^2 или n^3 раз, где n – количество кадров. Технология Enhanced Super Resolution Generative Adversarial Network (ESRGAN) показала свою эффективность в улучшении качества видео, что стало значительным достижением в этой сфере. Будущие улучшения архитектуры позволят добиться еще большей скорости и качества обработки [10].

Потенциальные приложения в будущем

В будущем ожидается, что технология GAN будет активно использоваться в медицине для улучшения качества снимков, что поможет врачам лучше интерпретировать результаты исследований по снимкам. Так же технология может стать важным инструментом в космической съемке, где для анализа космических объектов таких как планеты, звезды и целые галактика требуются высококачественные изображения [6].

4. Заключение

В сфере компьютерного зрения генеративно-сопоставительные сети (GAN) стремительно набирают популярность, особенно когда речь идет об улучшении разрешения изображений. Такие модели, как SRGAN и ESRGAN, действительно помогают сделать картинки четче и качественнее. Одной из основных проблем традиционных методов повышения разрешения является потеря мелких деталей, с которой GAN справляются гораздо лучше. Вместо простого добавления пикселей, эти сети восстанавливают текстуры и детали, которые раньше терялись. Это особенно полезно в таких областях, как реставрация старых фильмов, видеонаблюдение, спутниковая съемка и медицинская визуализация.

Технология генеративно-сопоставительных сетей стремительно набирает популярность в сфере улучшения качества изображений. Например, такие модели как SRGAN и ESRGAN, помогают сделать изображения сильно качественнее и четче. Основной проблемой повышения изображения на основе традиционных методов повышения разрешения является потеря мелких деталей. В свою очередь технология GAN гораздо лучше справляется с этой задачей. Вместо добавления пикселей, сети могут восстановить текстуры и детали, которые раньше могли быть утеряны. Однако использование GAN не обошлось без проблем. Во-первых, они требуют больших вычислительных ресурсов, а во-вторых, для их обучения нужны огромные наборы данных. Эти факторы усложняют внедрение GAN на практике, несмотря на их очевидные преимущества.

Перспективы развития этих технологий остаются многообещающими. Ожидается улучшения работы генеративно-сопоставительных сетей в реальном времени. Модели GAN могут адаптироваться к ещё более широкому кругу задач, суперразрешение может быть очень полезно во множестве областей от медицины до бытовых сфер.

Список используемых источников

1. Житников В.П., Шерыхалина Н.М., Федорова Г.И., Соколова А.А. Методика качественного улучшения результатов вычислительного эксперимента // Системная инженерия и информационные технологии. 2021. Т. 3, №1 (5). С. 58-64. [[Zhitnikov V.P., Sherykhalina N.M., Fedorova G.I., Sokolova A.A., "Methodology for qualitative improvement of the results of a computational experiment". System Engineering and Information Technologies, 2021, vol. 3, no 1 (5), pp. 58-64.]]
2. Ledig, C., Theis, L., Huszar, F. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — Vol. 3, № 6. — P. 4681–4690.
3. Радфорд, А., Метц, Л., Чинтала, С. Неподконтрольное обучение представлений с помощью глубоких сверточных GAN // arXiv preprint. — 2015. — № 3. — С. 15–28.
4. Wang, X., Yu, K. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks // Proceedings of the European Conference on Computer Vision (ECCV) Workshops. — 2018. — Vol. 2. — P. 63–79.
5. Goodfellow, I., Pouget-Abadie, J., Bengio, Y. Generative Adversarial Nets // Advances in Neural Information Processing Systems. — 2014. — Vol. 27. — P. 2672–2680.
6. Васильев, И.П. Применение нейронных сетей для улучшения изображений в спутниковой съемке //

- Труды Академии Наук. — 2020. — № 6. — С. 78–90.
7. Sherykhalina N.M., Saifullin R.O., Shaymardanova E.R. Multidimensional polynomial interpolation // Systems Engineering and Information Technologies. — 2023. — Vol. 5, № 4 (13). — P. 94–100.
 8. Киселев, А.В. Применение ESRGAN для восстановления старых фильмов // Журнал «Киноиндустрия». — 2020. — № 4. — С. 42–48.
 9. Харламов, С.А. Обзор методов суперразрешения изображений // Neural Networks Journal. — 2019. — № 2. — С. 55–62.
 10. Zhang, H., Goodfellow, I. Recent Advances in Generative Adversarial Networks: Theory and Practice // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2022. — Vol. 44, № 4. — P. 1234–1246.
 11. Kim, J., Park, S., Cho, Y. Real-Time Image Enhancement Using GANs in Autonomous Driving Systems // Journal of Autonomous Vehicles and Artificial Intelligence. — 2022. — Vol. 9, № 2. — P. 78–90.
 12. Yu, Z., Liu, Q., He, Y. GAN-based Super-Resolution for Satellite Imagery: A Comprehensive Review // Remote Sensing. — 2021. — Vol. 13, № 4. — P. 75–89.
 13. Zhang, X., Zhao, S., Zhang, Y. Smart City Implementation with GAN-based Video Enhancement: Enhancing Safety Through Better Resolution // Journal of Visual Communication and Image Representation. — 2022. — Vol. 86. — P. 22–37.
 14. Lim, S., Kang, H. GAN-TTS: A Generative Model for Time-Series Super-Resolution // IEEE Transactions on Neural Networks and Learning Systems. — 2023. — Vol. 34, № 2. — P. 312–329.
 15. Wu, J., Lu, Y., Zhang, S. Deep Learning Methods for Image Super-Resolution: A Survey // ACM Computing Surveys. — 2021. — Vol. 54, № 2. — P. 1–35.
 16. Wang, X., Gu, J., Dong, C., Loy, C.C. Deep Generative Networks for Real-Time Image Super-Resolution // Proceedings of the IEEE International Conference on Computer Vision (ICCV). — 2019. — P. 2040–2049.
 17. Park, S., Lee, K. Self-Supervised GAN-based Super-Resolution: Improving Image Quality with Limited Labeled Data // Pattern Recognition Letters. — 2022. — Vol. 159. — P. 87–96.
 18. Tian, C., Xu, Y., Zuo, W. Image Super-Resolution Using Lightweight GAN Networks // IEEE Transactions on Image Processing. — 2021. — Vol. 30. — P. 1983–1997.
 19. Shen, Z., Liu, X., Cao, J. Efficient GAN-based Models for Super-Resolution: Compression and Quantization Techniques // Neural Processing Letters. — 2022. — Vol. 54, № 1. — P. 35–52.
 20. Park, S., Kim, J. Optimized Super-Resolution Networks with Quantization and Pruning for Embedded Systems // Journal of Real-Time Image Processing. — 2021. — Vol. 18, № 3. — P. 410–425.
 21. Chen, Y., Lee, S., Chang, H. Self-Supervised Learning for Image Super-Resolution: A Survey of Recent Techniques // International Journal of Computer Vision. — 2021. — Vol. 129, № 2. — P. 457–478.
 22. Jo, J., Choi, W., Han, S. Reducing Artifacts in GAN-Generated Images Through Enhanced Loss Functions // Journal of Artificial Intelligence Research. — 2022. — Vol. 65, № 5. — P. 77–90.