

УДК 004.89

doi 10.54708/22259309\_2026\_13571

## ПРИМЕНЕНИЕ GAN В РЕАЛЬНОМ МИРЕ: КАК ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫЕ СЕТИ УЛУЧШАЮТ КАЧЕСТВО ИЗОБРАЖЕНИЙ

Е. В. САВИН<sup>1</sup>

<sup>1</sup>n778hr@mail.ru

<sup>1</sup>ФГБОУ ВО «Уфимский университет науки и технологий» (УУНИТ)

**Аннотация.** В работе рассмотрены архитектура и принципы функционирования генеративно-состязательных сетей (GAN), а также примеры их применения в обработке изображений. Приведены конкретные технологии — MedSRGAN, Retinal-ESRGAN и DragGAN, каждая из которых решает уникальные задачи, связанные с повышением разрешения и качеством изображений. Показано, как GAN позволяют сохранять диагностически значимую информацию и обеспечивают возможности интерактивного редактирования. Работа демонстрирует, что GAN представляют собой эффективный инструмент в современной медицинской визуализации и компьютерной графике.

**Ключевые слова:** генеративно-состязательные сети; GAN; глубокое обучение; улучшение качества изображений; MedSRGAN; Retinal-ESRGAN; DragGAN; суперразрешение; медицинская визуализация; компьютерное зрение; нейронные сети; интерактивное редактирование изображений; КТ-изображения; офтальмологические снимки.

### ВВЕДЕНИЕ

В представленной работе рассмотрены принципы работы генеративно-состязательных сетей (GAN), а также их применение в области обработки изображений. Данный класс нейросетевых моделей используется для генерации реалистичных изображений, повышения их разрешения и редактирования визуального контента. Основу архитектуры GAN составляют две нейросети – генератор и дискриминатор, которые обучаются в состязательной форме. Генератор создает изображения, стремясь сделать их максимально приближенными к реальным, в то время как дискриминатор пытается отличить подделку от оригинала.

GAN находят широкое применение в различных областях, включая медицину, где высокое качество изображений критически важно для точной диагностики. Наиболее яркими примерами служат модели MedSRGAN и Retinal-ESRGAN, разработанные для повышения качества медицинских изображений, а также DragGAN – интерактивный инструмент редактирования, обеспечивающий точечную манипуляцию с изображениями. Рассматриваемые технологии демонстрируют эффективность GAN в решении прикладных задач, связанных с визуальными данными.

### ПРИНЦИП РАБОТЫ GAN

Эта система имеет два главных компонента: генератор G – нейронная сеть, которая начинает работать от шума с нормальным распределением. Генератор создает данные, пытаясь максимально приблизить их к заданному реальному распределению. Дискриминатор D – вторичная нейронная сеть получает на вход как настоящие, так и сгенерированные данные. Задача дискриминатора – определить, где подделка, а где оригинал.

Обучение генеративно-сопоставительных сетей представляет собой минимаксную игру с функцией оптимизации [1], где две нейронные сети соревнуются друг с другом: одна пытается создавать всё более реалистичные данные, а другая стремится улучшить свою способность отличать искусственные данные от настоящих.

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

$x$  – реальное изображение из распределения  $P_{data}$

$z$  – случайный шум из распределения  $P_z$

Генератор обучается так, чтобы максимизировать вероятность ошибки дискриминатора, он же в свою очередь стремится минимизировать вероятность ошибки классификации. Это можно представить как игру с нулевой суммой между двумя агентами, где победа одного является проигрышем другого [2].

### MEDSRGAN

Синтетические данные стали важной частью медицинской визуализации. В диагностике качество изображений имеет большое значение для правильного выявления заболеваний и планирования лечения. Генеративно-сопоставительная архитектура MedSRGAN представляет значительный прорыв в области медицинской визуализации. Данная технология решает критическую проблему баланса между качеством изображения и дозовой нагрузкой на пациента.

Основу MedSRGAN составляет генератор RWMAN (Residual Whole Map Attention Network). В отличие от стандартных алгоритмов интерполяции, которые просто увеличивают разрешение, RWMAN умеет различать клиническую важность разных частей изображения. Анатомические структуры получают приоритет обработки в зависимости от их диагностической ценности.

Дискриминатор работает как строгий контролер качества – он сравнивает исходные и восстановленные изображения. Комплексная система функций потерь гарантирует сохранение диагностически значимой информации при одновременном достижении фотореалистичности результата. Валидация проводилась на репрезентативных выборках: 242 КТ-исследования органов грудной клетки и 110 МРТ-сканов головного мозга [4]. Радиологи оценивали результаты и подтвердили: текстуры и анатомические детали в обработанных изображениях сохраняются.

Технология демонстрирует особую эффективность при работе с протоколами КТ и МРТ-исследованиями на аппаратах с низкой напряженностью магнитного поля. Врачи могут снижать лучевую нагрузку на пациентов, не жертвуя качеством диагностики.

Отдельное исследование по выявлению легочных узелков показало, что реконструированные изображения не уступают оригинальным данным в задачах патологического скрининга. Это подтверждает клиническую применимость технологии в практической диагностике. Авторы исследования планируют расширение области применения методики, что может существенно изменить подходы к медицинской визуализации в ближайшем будущем. Такие технологии, возможно, могут кардинально изменить доступность качественной диагностики, особенно в местах с не самым современным оборудованием.

### RETINAL-ESRGAN

Суперразрешение медицинских изображений стало важным инструментом в диагностике, особенно в офтальмологии, где четкие снимки сетчатки нужны для раннего выявления проблем [4]. Новая разработка Retinal-ESRGAN – это модель генеративно-сопоставительных сетей, которая улучшает разрешение снимков сетчатки с меньшими затратами на оборудование и время. Retinal-ESRGAN улучшает оригинальную модель

ESRGAN и добавляет функции для повышения качества. Здесь комбинируются два подхода: один фокусируется на точности на уровне пикселей, а другой – на улучшении визуального качества через соперничество в обучении. Разработчики нашли способ разделить обучение на маленькие части, чтобы облегчить нагрузку на ресурсы [5].

Результаты тестирования Retinal-ESRGAN заставляют пересмотреть устоявшиеся представления о возможностях реконструкции медицинских изображений. Превосходство над SRGAN и ESRGAN очевидно не только в метриках PSNR и SSIM, но и в том аспекте, который действительно имеет значение для практикующих врачей – способности различать тонкие патологические изменения сетчатки [6]. Парадоксально, но снижение вычислительных требований при росте качества ставит вопрос о том, почему подобные решения не появились раньше.

Разработчики честно указывают на недостатки – усложнение архитектуры и необходимость тонкой настройки параметров. Тем не менее для офтальмологии это серьезный шаг вперед, особенно в плане ранней диагностики заболеваний сетчатки глаза [5].

### **DRAGGAN: РЕВОЛЮЦИЯ В ИНТЕРАКТИВНОМ РЕДАКТИРОВАНИИ СИНТЕТИЧЕСКИХ ИЗОБРАЖЕНИЙ**

Исследователи из Института Макса Планка создали DragGAN – инструмент для интерактивного редактирования GAN-изображений через простое перетаскивание элементов [7].

Модуль направленного движения признаков анализирует латентное пространство генератора, он в свою очередь локализует область влияния выбранной контрольной точки. Система конструирует специализированную функцию потерь, которая векторизует движение признаков к заданной цели, при этом применяя однонаправленное ограничение для предотвращения геометрических деформаций. Параллельный механизм трекинга обеспечивает непрерывное отслеживание точек в процессе трансформации. После каждой итерации алгоритм реидентифицирует позиции контрольных точек через поиск семантически эквивалентных признаков в обновленном изображении. Именно такая стратегия обеспечивает интерактивность процесса в режиме реального времени. Итерационный цикл завершается при достижении целевых координат. Можно сказать, что пользователь может наблюдать вживую целое перемещение по латентному пространству и все итерации при смещении координат изображения.

Функциональные возможности DragGAN характеризуются исключительной масштабируемостью применения. Технология демонстрирует стабильную производительность при работе с гетерогенными классами объектов: портретными изображениями, зоологическими объектами, автомобильной техникой и природными пейзажами. Система предоставляет пользователю точный контроль над множественными атрибутами изображения – от корректировки позиции и мимических характеристик до селективного редактирования выделенных регионов с сохранением неизменности остальных областей [8].

Особую ценность представляет способность системы интегрировать реальные фотографии через их проекцию в латентное пространство GAN-архитектуры.

Сравнения с существующими методами показывают преимущество DragGAN. Особенно важно, что технология может работать с реальными фотографиями через их проекцию в GAN-пространство. Результат – это довольно сложные изменения простыми действиями при сохранении реалистичности изображения [7].

### **BRAIN MRI SUPER-RESOLUTION**

Brain MRI Super-Resolution using 3D GAN – это инновационная технология для повышения разрешения МРТ-изображений головного мозга, разработанная исследователями в области медицинской визуализации. С помощью этого метода

специалисты могут получать высококачественные детализированные снимки из изображений низкого разрешения, используя возможности трёхмерных генеративно-сопоставительных сетей [9].

Технология основывается на двух ключевых компонентах. Первый – это усовершенствованный дискриминатор с функцией потерь наименьших квадратов. Вторая часть системы – это трёхмерная архитектура генератора, адаптированная от классической SRGAN. Система использует 3D-свёртки для извлечения объёмной информации из МРТ-данных, создавая комбинированную функцию потерь на основе сопоставительной потери и контентной потери MSE. Это позволяет сохранять анатомическую точность восстанавливаемых структур мозга и учитывать пространственные связи между срезами.

Brain MRI Super-Resolution using 3D GAN превосходит традиционные методы интерполяции по ключевым метрикам качества. Технология позволяет сохранять анатомическую целостность, контролировать восстановление мелких деталей и значительно сокращать время получения диагностических изображений. Этот инструмент представляет собой прорыв в медицинской визуализации, позволяя получать детализированные изображения мозга без увеличения времени сканирования, сохраняя при этом диагностическую ценность данных.

### ЗАКЛЮЧЕНИЕ

В представленной статье был проведен обзор генеративно-сопоставительных сетей (GAN) и их применения в области улучшения качества изображений. Рассмотрены четыре ключевые технологии: MedSRGAN для улучшения медицинских изображений, Retinal-ESRGAN для повышения качества снимков сетчатки глаза и DragGAN для интерактивного редактирования изображений; Brain MRI Super-Resolution using 3D GAN – это нейронная сеть, предназначенная для повышения чёткости и детализации МРТ-снимков головного мозга из изображений низкого разрешения. Анализ этих технологий показывает, что современные генеративно-сопоставительные сети совершили значительный прогресс от основополагающей работы И. Гудфеллоя и его коллег до специализированных решений, адаптированных под конкретные задачи. Особого внимания заслуживает их практическое применение в медицине, где повышение качества изображений критически важно для точной диагностики заболеваний. MedSRGAN демонстрирует способность сохранять диагностически важные детали при улучшении КТ и МРТ-изображений, Retinal-ESRGAN достигает высоких показателей качества для офтальмологических снимков, а DragGAN предлагает интуитивный интерфейс для манипуляций с изображениями.

### СПИСОК ЛИТЕРАТУРЫ

1. **Goodfellow, I.** Generative Adversarial Nets / I. Goodfellow, J. Pouget-Abadie, M. Mirza [и др.] // *Advances in Neural Information Processing Systems*. – 2014. – Vol. 27. – P. 2672–2680.
2. **Горбунов, Е. В.** Применение генеративно-сопоставительных сетей для синтеза изображений лиц с учетом возрастных изменений / Е. В. Горбунов, С. А. Козырев // *Труды конф. «Интеллектуальные системы»*. – 2016. – С. 48–53.
3. **Gu, Yuchun.** MedSRGAN: Medical Image Super-Resolution Using Generative Adversarial Networks / Y. Gu, Z. Zeng, H. Chen [и др.] // *Scientific Reports*. – 2022. – Vol. 12. – Article number: 9533. – 12 p. – DOI: 10.1038/s41598-022-13658-4.
4. **Rieck, B.** The Impact of GANs on Healthcare Data Generation / B. Rieck, K. Gajos, S. Albayrak // *Journal of Medical Imaging and Health Informatics*. – 2020. – Vol. 10. – № 3. – P. 1221–1229.
5. **Deepthi, K.** Retinal-ESRGAN: A Hybrid GAN Model Approach for Retinal Image Super-Resolution / K. Deepthi, K. Aditya Shastry // *IEEE Access*. – 2023. – Vol. 11. – P. 61405–61418.
6. **Kamran, S. A.** RV-GAN: Segmenting Retinal Vascular Structure... / S. A. Kamran, K. F. Hossain, A. Tavakkoli [и др.] // *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. – Cham: Springer, 2021.
7. **Pan, X.** Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold / X. Pan, A. Tewari, T. Leimkühler [и др.] // *Proc. of the ACM SIGGRAPH 2023 Conference*. – New York: ACM, 2023.
8. **Liu, C.** A Technical Deep Dive into Drag Your GAN (DragGAN) [Electronic resource] / C. Liu. – 2023.
9. **Chen, Y.** Brain MRI Super-Resolution Using 3D Generative Adversarial Networks / Y. Chen, F. Shi, A.G. Christodoulou [и др.] // *NeuroImage*. – 2018. – Vol. 174. – P. 550–562. – DOI: 10.1016/j.neuroimage.2018.03.038.

**ОБ АВТОРАХ**

**САВИН Егор Вячеславович**, студент, Математическое обеспечение и администрирование информационных систем, Институт информатики, математики и робототехники УУНиТ.

**METADATA**

**Title:** REAL-WORLD APPLICATION OF GAN: HOW GENERATIVE ADVERSARIAL NETWORKS IMPROVE IMAGE QUALITY

**Author:** E.V. Savin<sup>1</sup>

**Affiliation:**

<sup>1</sup> Ufa University of Science and Technology (UUST), Russia.

**Email:** [n778hr@mail.ru](mailto:n778hr@mail.ru)

**Language:** Russian.

**Source:** Molodezhnyj Vestnik UGATU (scientific journal of Ufa University of Science and Technology), no. 1 (35), pp. 71-75, 2026. ISSN 2225-9309 (Print).

**Abstract:** The paper considers the architecture and principles of operation of generative adversarial networks (GANS), as well as examples of their application in image processing. Specific technologies are given — MedSRGAN, Retinal-ESRGAN and DragGAN, each of which solves unique tasks related to increasing resolution and image quality. It is shown how GANS allow you to save diagnostically significant information and provide interactive editing capabilities. The work demonstrates that GANS are an effective tool in modern medical imaging and computer graphics.

**Key words:** Generative-adversarial networks, GAN, deep learning, image quality improvement, MedSRGAN, Retinal-ESRGAN, DragGAN; super resolution, medical imaging, computer vision, neural networks, interactive image editing, CT images, ophthalmic images.

**About authors:**

**SAVIN Egor Vyacheslavovich**, a student of the Mathematical support and administration of information systems Institute of Computer Science, Mathematics and Robotics, UUST