

DOI: 10.15514/ISPRAS-2026-38(2)-1



## Применение состязательных атак для локализованной генерации артефактов суперразрешения

<sup>1</sup> К.В. Мальшев, ORCID: 0009-0003-3698-0143 <kirill.malyshev@graphics.cs.msu.ru>

<sup>1,2</sup> И.А. Молодецких, ORCID: 0000-0002-8294-0770 <ivan.molodetskikh@graphics.cs.msu.ru>

<sup>1,2,3,4</sup> Д.С. Ватолин, ORCID: 0000-0002-8893-9340 <dm.vatolin@ispras.ru>

<sup>1</sup> Московский государственный университет имени М.В. Ломоносова, Россия, 119991, Москва, Ленинские горы, д. 1.

<sup>2</sup> Центр искусственного интеллекта МГУ.

<sup>3</sup> Институт искусственного интеллекта МГУ.

<sup>4</sup> Институт системного программирования им. В.П. Иванникова РАН, Россия, 109004, г. Москва, ул. А. Солженицына, д. 25.

**Аннотация.** Задача суперразрешения изображений, решаемая с помощью глубоких нейросетей, особенно генеративно-состязательных моделей, сталкивается с проблемой появления визуальных артефактов. Эти искажения ухудшают качество результата, а их автоматическое обнаружение затруднено из-за отсутствия крупных размеченных наборов данных. Целью данной работы является разработка автоматизированного метода создания таких наборов данных для обучения и оценки моделей детекции артефактов. Предлагаемый метод использует подход состязательных атак для целенаправленного создания артефактов в выходных изображениях моделей суперразрешения. В основе метода лежит модификация итеративной атаки быстрого знака градиента. Ключевое нововведение заключается в модификации функции потерь, которая позволяет максимизировать искажения в заданной области изображения, определяемой бинарной маской, и одновременно минимизировать их в остальных частях. Это обеспечивает генерацию локализованных артефактов, имитирующих естественные дефекты. Для проверки метода был создан набор данных, содержащий более 2 тысяч примеров. Результаты экспериментов показали, что предложенный набор данных обладает качественной разметкой. Методы детекции продемонстрировали на нем значение IoU более 0.7, что существенно выше результатов, достигнутых на существующих наборах данных. Разработанный метод позволяет эффективно создавать масштабируемые и качественно размеченные наборы данных. Также был разработан нейросетевой метод, который показывает более качественные результаты по сравнению с базовым подходом. Это открывает возможности для разработки более надежных методов суперразрешения, их последующей постобработки и создания эффективных детекторов артефактов.

**Ключевые слова:** суперразрешение изображений; артефакты изображений; генеративно-состязательные сети; состязательные атаки; метод быстрого знака градиента FGSM; семантическая сегментация; создание наборов данных; детекция артефактов.

**Для цитирования:** Мальшев К.В., Молодецких И.А., Ватолин Д.С. Применение состязательных атак для локализованной генерации артефактов суперразрешения. Труды ИСП РАН, том 38, вып. 2, 2026 г., стр. 7–20. DOI: 10.15514/ISPRAS-2026-38(2)-1.

## Using Adversarial Attacks for Localized Generation of Super-Resolution Artifacts

<sup>1</sup> K.V. Malyshev, ORCID: 0009-0003-3698-0143 <kirill.malyshev@graphics.cs.msu.ru>

<sup>1,2</sup> I.A. Molodetskikh, ORCID: 0000-0002-8294-0770 <ivan.molodetskikh@graphics.cs.msu.ru>

<sup>1,2,3,4</sup> D.S. Vatolin, ORCID: 0000-0002-8893-9340 <dm.vatolin@ispras.ru>

<sup>1</sup> Lomonosov Moscow State University, GSP-1, Leninskie Gory, Moscow, 119991, Russia.

<sup>2</sup> MSU AI Center.

<sup>3</sup> MSU Institute for Artificial Intelligence.

<sup>4</sup> Ivannikov Institute for System Programming of the Russian Academy of Sciences, 25, Alexander Solzhenitsyn st., Moscow, 109004, Russia.

**Abstract.** The task of image super-resolution, addressed using deep neural networks, particularly generative adversarial models, faces the problem of visual artifacts. These distortions degrade the result quality, and their automatic detection is challenging due to the lack of large-scale labeled datasets. This work aims to develop an automated method for creating such datasets to train and evaluate artifact detection models. The proposed method utilizes an adversarial attack approach to deliberately create artifacts in the output images of super-resolution models. The core of the method is a modification of the Iterative Fast Gradient Sign Method. The key innovation lies in the modified loss function, which maximizes distortions in a specified image area, defined by a binary mask, while simultaneously minimizing them in the remaining parts. This enables the generation of localized artifacts that mimic natural defects. To validate the method, a dataset containing over 2000 examples has been created. Experimental results confirmed that the proposed dataset possesses high-quality annotations. Detection methods demonstrated an IoU value exceeding 0.7 on it, which is substantially higher than results achieved on existing datasets. The developed method allows for the efficient creation of scalable and high-quality labeled datasets. A neural network method was also developed, which shows better results compared to the baseline method. This opens up opportunities for developing more robust super-resolution methods, their subsequent post-processing, and creating effective artifact detectors.

**Keywords:** image super-resolution; image artifacts; generative adversarial networks; adversarial attacks; FGSM; semantic segmentation; dataset creation; artifact detection.

**For citation:** Malyshev K.V., Molodetskikh I.A., Vatolin D.S. Using Adversarial Attacks for Localized Generation of Super-Resolution Artifacts. Trudy ISP RAN/Proc. ISP RAS, vol. 38, issue 2, 2026, pp. 7-20 (in Russian). DOI: 10.15514/ISPRAS-2026-38(2)-1.

### 1. Введение

Задача суперразрешения изображений заключается в повышении разрешения и детализации изображения, как правило, с увеличением его линейных размеров в заданное количество раз. В последние годы данная задача решается с использованием глубоких нейросетевых моделей.

Типичным коэффициентом масштабирования является увеличение длины и ширины изображения в 4 раза, что соответствует увеличению общего количества пикселей в 16 раз. Таким образом, для каждого исходного пикселя модели необходимо реконструировать 15 новых пикселей.

Следует отметить, что различные подходы к обучению моделей демонстрируют различное качество реконструкции. Модели, обученные с применением функций потерь, основанных на метриках пиксельной схожести, таких как среднеквадратичная ошибка (MSE) или средняя абсолютная ошибка (MAE), часто обеспечивают высокий показатель пиковое отношение сигнала к шуму (PSNR), но могут страдать от недостаточной чёткости и недостаточно точного восстановления высокочастотных деталей, что приводит к излишне сглаженным

результатам. В свою очередь, модели, обученные с применением состязательных функций потерь (GAN), способны генерировать результаты с более высокой визуальной правдоподобностью и чёткостью деталей. Однако данный подход может индуцировать возникновение артефактов и искажений текстуры, не присутствовавших в исходном изображении (рис. 1).



Рис. 1. Пример появления текстурного артефакта в результате применения метода суперразрешения Real-ESRGAN [1] (справа) в сравнении с оригинальным изображением высокого разрешения (слева).

Fig. 1. An example of the appearance of a texture artifact as a result of applying the Real-ESRGAN [1] super-resolution method (right) compared to the original high-resolution image (left).

В связи с указанной склонностью моделей, обученных на состязательных функциях потерь, к генерации артефактов и искажений, актуальной становится задача обнаружения и локализации подобных дефектов.

Решение этой задачи открывает возможности для нескольких практических применений. Во-первых, для постобработки и коррекции итогового изображения. Выделив области с артефактами, можно предпринять меры по их устранению, например, локально регенерировать эти фрагменты с использованием альтернативных методов. Это может быть выполнено с помощью другой, более консервативной модели суперразрешения (например, оптимизированной на MSE) или даже с применением классических методов интерполяции, таких как бикубическая, что позволяет гарантировать отсутствие грубых искажений в критических участках изображения.

Во-вторых, карта локализованных артефактов может служить ценным источником информации для улучшения и дообучения самих методов суперразрешения. Включение данных об артефактах в функцию потерь или использование их для отбора "сложных" областей во время обучения позволяет направить оптимизацию модели не только на повышение объективного качества, но и на подавление нежелательных искажений при генерации текстур. Таким образом, задача детекции артефактов становится ключевым элементом в разработке более надежных и точных алгоритмов повышения разрешения.

Однако разработка моделей для детекции подобных артефактов сама по себе сопряжена с существенной методологической проблемой – отсутствием обширных и качественно размеченных наборов данных. Для эффективного обучения таких моделей требуется набор данных, содержащий парные примеры: изображения, полученные в результате суперразрешения, и соответствующие им пиксельные маски, точно выделяющие области с артефактами.

Создание такого набора данных является сложной и ресурсоемкой задачей. Основная трудность заключается в необходимости проведения субъективной экспертной оценки для

идентификации артефактов, поскольку их характер может быть разнообразным и не всегда однозначно определяемым алгоритмически. Более того, процесс точной разметки областей (сегментации) артефактов требует привлечения квалифицированных разметчиков, является крайне трудоемким и, как следствие, финансово затратным. Отсутствие крупномасштабных наборов данных с подобной разметкой в настоящее время сдерживает развитие надежных методов контроля качества в задачах суперразрешения.

В данной работе предлагается автоматизированный метод создания наборов данных артефактов, возникающих в моделях суперразрешения, основанный на методах состязательных атак. Такой метод позволяет гибко управлять интенсивностью артефактов и областью их локализации, а также генерировать наборы данных разного масштаба для широкого спектра моделей суперразрешения.

Еще одним результатом работы стал автоматизированно созданный набор данных, включающий 2101 пример изображений с артефактами суперразрешения. Проведенная оценка показала, что современные методы детекции демонстрируют повышенную эффективность на предложенном наборе данных и по показателю IoU превосходят результаты, полученные на наборе данных с ручной разметкой DeSRA [2].

Для дополнительного подтверждения пригодности набора данных в прикладных задачах был разработан нейросетевой метод детекции артефактов суперразрешения, при обучении которого использовался предлагаемый набор данных. Обученная модель показала более высокие объективные результаты для двух из трёх методов суперразрешения по сравнению с базовым методом DeSRA на вручную размеченном наборе данных DeSRA.

## 2. Обзор литературы

### 2.1 Наборы данных

Составление набора изображений с размеченными текстурными артефактами является трудоемким процессом, который требует больших временных и финансовых затрат. Разметка осложняется тем, что необходимо получить не только маску артефакта, но и оценить его «качество». В дополнение, далеко не все изображения в результате работы того или иного метода суперразрешения будут содержать артефакты.

К тому же данные, размеченные для одного метода суперразрешения, могут терять свою применимость при работе с другими методами. Иными словами, разные методы суперразрешения могут генерировать артефакты разного характера и силы.

Тем не менее, существуют работы, в которых предлагаются наборы изображений с размеченными артефактами. Так, авторы метода детекции артефактов методов суперразрешения DeSRA [2] составили и вручную разместили набор данных из 600 изображений. Набор содержит по 200 изображений для 3 методов суперразрешения. Авторы использовали набор данных для дообучения соответствующих методов суперразрешения с целью сделать их более устойчивыми к генерации артефактов.

В то же время существуют наборы изображений с артефактами и для других задач. Авторы метода PAL4Inpaint [3] составили набор из 4795 изображений, где артефакт возникает в результате работы метода закрасивания (inpainting). В свою очередь авторы PAL4VST [4] собрали набор из 10168 примеров артефактов для различных задач синтеза изображений.

### 2.1 Методы детекции артефактов

В работе [5] предлагается метод LDL для оценки артефактов на основе анализа локальной дисперсии разности оригинального изображения высокого разрешения и полученного на выходе метода суперразрешения. Авторы используют такой подход внутри функции потерь

при обучении метода суперразрешения, чтобы обеспечить устойчивость к генерации артефактов.

Авторы DeSRA [2] развивают идею LDL и предлагают более универсальную относительную оценку наличия артефактов для локальных областей. Такой подход может использоваться для дообучения существующих методов суперразрешения. Также авторы модифицируют показатели объективной оценки Precision и Recall, адаптируя их к задаче детекции артефактов методов суперразрешения.

Наряду с описанными выше методами, существует большое количество методов оценки качества изображений, которые учитывают текстурные искажения, но не предназначены для обнаружения артефактов. Однако их побочное применение и последующая бинаризация могут работать в качестве метода детекции артефактов. Например, такими методами являются SSIM [6] и DISTS [7].

### 3. Предлагаемый метод

Состязательными атаками называют целенаправленные воздействия на машинное обучение, при которых в исходные данные вносятся малозаметные для человека возмущения, что приводит к ошибочным предсказаниям модели. Состязательные атаки часто применяют для искажения результатов моделей классификации изображений или объективной оценки качества. Такой подход можно применить и к моделям суперразрешения, чтобы модифицировать входное изображение малого разрешения так, что на выходе будет получено изображение высокого разрешения, которое содержит артефакты.

Для этого необходимо определить функцию потерь, которую атака будет максимизировать. В самом базовом случае это будет различие  $L(SR(I_{lr} + \delta^0), I_{hr})$  между изображением  $SR(I_{lr} + \delta^0)$ , полученным на выходе модели суперразрешения  $SR$ , и оригинальным изображением высокого разрешения  $I_{hr}$ . Слагаемое  $\delta^0$  является начальным возмущением, которое, как правило, нулевое,  $\delta^1$  – атакующее возмущение. Различие между изображениями может быть выражено как классическими функциями потерь (MAE, MSE), так и более сложными эталонными метриками оценки объективного качества (LPIPS [8], DISTS и др.).

В данной работе рассматривается широко используемая состязательная атака методом быстрого знака градиента FGSM (Fast Gradient Sign Method) [9]:

$$\delta^1 = \delta^0 + \alpha * \text{sign}(\nabla_{\delta^0} L(SR(I_{lr} + \delta^0), I_{hr})) \quad (1)$$

и её итеративная модификация I-FGSM (Iterative Fast Gradient Sign Method), в которой метод FGSM применяется итеративно несколько раз для усиления атакующего эффекта.

Применение такой состязательной атаки позволяет получить возмущение, в результате добавления которого к входному изображению низкого разрешения метод суперразрешения сгенерирует большое количество артефактов в выходном изображении. Такие изображения (рис. 2) уже могут использоваться как примеры некорректной работы методов суперразрешения, например для построения метода защиты от состязательных атак или для дообучения исходного метода суперразрешения. Однако это не будет похоже на естественные примеры некорректной работы методов суперразрешения, в которых искажение текстур обычно локализовано в некоторой области, которая, как правило, примерно соответствует границам некоторого объекта на изображении.

Функция потерь может быть модифицирована так, чтобы состязательная атака максимизировала разницу только в некоторой избранной области

$$\delta^1 = \delta^0 + \alpha * \text{sign}(\nabla_{\delta^0} L(SR(I_{lr} + \delta^0) \odot \mathbf{M}, I_{hr} \odot \mathbf{M})) \quad (2)$$

Здесь  $\mathbf{M}$  – бинарная маска (рис. 3), которая выделяет объект (или область) для атаки, а  $\odot$  – операция поэлементного умножения.

Такая модификация позволяет сосредоточить атаку на искажении в маскированной области и существенно снизить влияние атакующего возмущения на области изображения (рис. 4), не выделенные маской. Однако поскольку такая функция потерь не учитывает влияние возмущения на пиксели вне маски, то, максимизируя искажение в выделенных областях, атака вносит некоторую долю визуально заметных искажений также и в остальные области, не выделенные маской. Это может быть связано с тем, что модели суперразрешения в той или иной степени учитывают глобальный контекст и максимизация функции потерь в одной области тянет за собой, хоть и в меньшей степени, подобные искажения в других областях.

Чтобы защитить области вне маски от искажающего действия состязательного возмущения, функция потерь может быть модифицирована так, чтобы одновременно максимизировать различие в маскированных областях и минимизировать его в остальных пикселях:

$$\delta^1 = \delta^0 + \alpha * \text{sign} \left( \nabla_{\delta^0} \left( L(SR(I_{lr} + \delta^0) \odot \mathbf{M}, I_{hr} \odot \mathbf{M}) - L(SR(I_{lr} + \delta^0) \odot (\mathbf{1} - \mathbf{M}), I_{hr} \odot (\mathbf{1} - \mathbf{M})) \right) \right) \quad (3)$$

Отметим, что такой подход не гарантирует, что в немаскированных областях не будет никаких изменений по сравнению с оригинальным выходом метода суперразрешения.

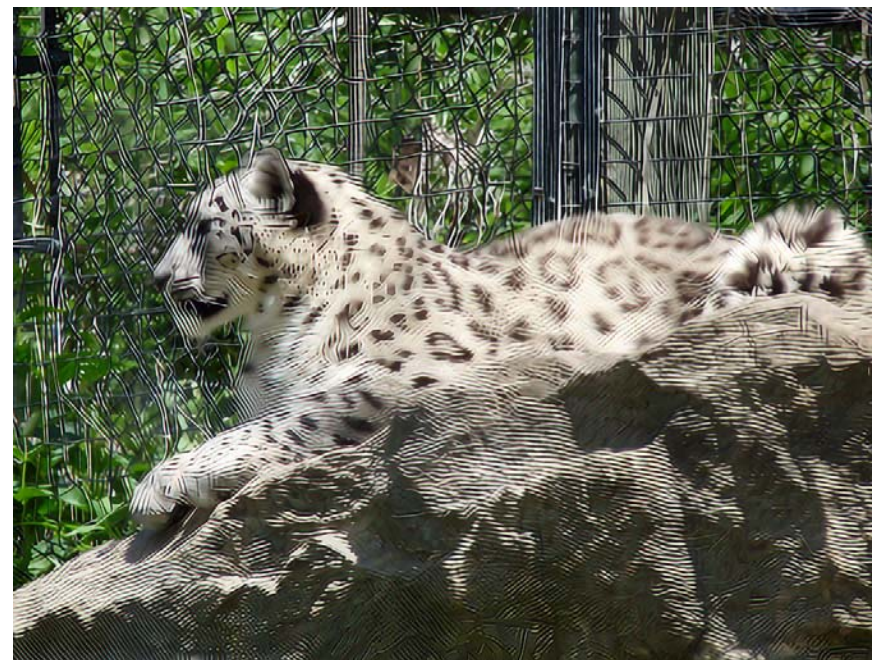


Рис. 2. Пример атаки I-FGSM на всю область результата работы метода суперразрешения.  
Fig. 2. An example of an I-FGSM attack on the entire super-resolution method result.



Рис. 3. Маска для применения состязательной атаки.  
Fig. 3. Mask for applying adversarial attack.



Рис. 4. Пример атаки I-FGSM на маскированную область результата работы метода суперразрешения.  
Fig. 4. An example of an I-FGSM attack on the masked area of the super-resolution method result.

Попиксельное сравнение показывает, что различия обычно есть во всех областях, однако они не носят характер текстурных искажений. Такой модификацией функции потерь можно получить метод состязательной атаки, который стабильно генерирует возмущения для получения искажений в выбранной области (рис. 5).

В качестве базового метода могут быть использованы и другие виды градиентных состязательных атак, не только FGSM.

#### 4. Эксперименты

Для проверки качества описанного подхода был создан набор данных с артефактами на основе метода суперразрешения DRCT [10] с применением состязательной атаки I-FGSM. Использовался подход с атакой на область по маске с защитой немаскированных пикселей

(формула 3). Атака I-FGSM применялась на протяжении пяти итераций с максимальной амплитудой возмущения  $\frac{5}{255}$  и коэффициентом возмущения  $\alpha = \frac{1}{255}$ . В качестве функции потерь использовалось среднее квадратичное отклонение (MSE).

Всего данные содержат 2101 изображение и являются подвыборкой набора Open Images [11]. Изображения в подвыборке имеют разрешение  $1024 \times 768$ , что удобно для экспериментов, так как применение градиентной состязательной атаки требует дополнительной видеопамати.

Выбор областей и получение масок для атаки производился автоматизированным образом с применением метода семантической сегментации SAN [12]. Атакованная область соответствует наиболее представленному классу по количеству пикселей на изображении. Эти же маски используются как эталонные и участвуют в объективной оценке. Отметим, что некоторые классы объектов, которые соответствуют фоновым или малозначимым областям и слабо подвергаются заметным текстурным искажениям (трава, небо и тому подобные), выбирались с пониженными приоритетами.



Рис. 5. Пример атаки I-FGSM на маскированную область результата работы метода суперразрешения с защитой немаскированной области.  
Fig. 5. An example of an I-FGSM attack on the masked area of the super-resolution method result with protection of the unmasked area.

Для созданного набора данных была произведена объективная оценка по методологии авторов метода DeSRA. Высокие объективные показатели позволяют сделать вывод, что собранный набор данных действительно содержит правильно размеченные артефакты, и методы детекции к ним применимы. При этом низкие показатели могут быть следствием низкого качества работы методов детекции артефактов. На рис. 6-9 приведены графики показателей IoU, точности (Precision) и полноты (Recall) для нескольких методов детекции артефактов.

Для получения результатов была использована версия SSIM, которая конструирует тепловую карту. Метод DISTS применялся по блокам для блоков  $16 \times 16$ .

Возможная немонотонность показателя полноты связана с тем, что методология DeSRA вводит показатели точности и полноты, которые вычисляются не обычным образом попиксельно, а по связным областям детекции по всему набору данных. Таким образом, уменьшение порогового значения может приводить к уменьшению количества детектированных областей и, как следствие, уменьшению показателя полноты.

Графики на рис. 6-9 показывают, что в лучшем случае все методы показывают значение IoU не менее 0.7, что является высоким показателем и заметно выше результатов авторов метода DeSRA на наборе данных из их работы при любых пороговых значениях. Это может свидетельствовать о низком качестве разметки набора данных из работы по DeSRA.

Показатели точности для всех методов монотонно возрастают и быстро достигают значения 1, что говорит о высокой корреляции сырых результатов работы методов с областями артефактов в наборе данных. Показатели полноты для методов SSIM и DISTs показывают значения, близкие к 1, для низких порогов, что говорит о том, что эти методы выделяют области артефактов почти полностью правильно, однако имеют и много ложноположительных результатов. Об этом свидетельствует относительно низкий показатель точности.

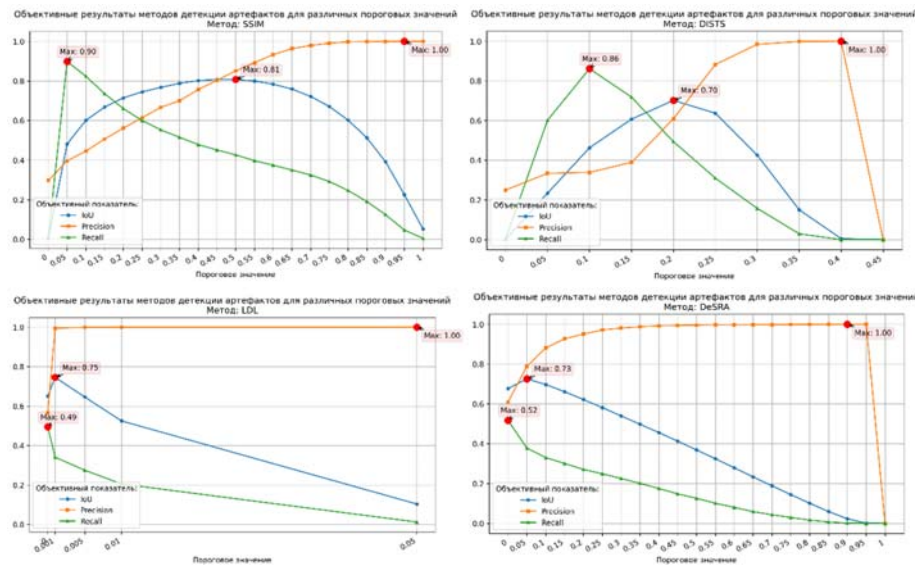


Рис. 6-9. Графики объективных показателей IoU, точности и полноты для различных методов детекции артефактов.

Fig. 6-9. Plots of IoU, Precision and Recall metrics for various artifact detection methods.

Низкая полнота для других методов и порогов может свидетельствовать о низкой чувствительности методов к некоторым артефактам или о том, что есть размеченные области, в которых артефакт слабо выражен. Однако лучший показатель  $\text{IoU} \geq 0.7$  говорит о том, что есть компромиссный порог для каждого метода, при котором он показывает высокое качество на предлагаемом наборе данных.

В табл. 1 представлено сравнение по показателю IoU предлагаемого набора данных с набором из работы DeSRA для метода суперразрешения Real-ESRGAN. Пороговые значения  $t \in \{0.01, 0.005, 0.001\}$  для метода LDL использованы, равные тем, что применяются в работе DeSRA. Для метода DeSRA выбраны следующие пороги  $t$ :

- $t = 0.3$  – стандартный порог, выбранный авторами DeSRA по наибольшему значению произведения точности на полноту;
- $t = 0.2$  – порог, который на наборе данных DeSRA показывает лучший результат;
- $t = 0$  – порог, полученный для предлагаемого набора данных по наибольшему значению произведения точности на полноту.

Следует отметить, что пороговое значение, равное 0, здесь не является тривиальным, так как бинаризация происходит по строгому неравенству.

Таким образом, оба рассмотренных метода показывают более хорошие результаты по IoU для предлагаемого набора данных по всем выбранным пороговым значениям. Это свидетельствует о более высококачественной разметке набора данных.

Чтобы подтвердить практическую применимость описанного подхода, на основе тепловых карт метода DeSRA была обучена модель архитектуры U-Net (рис. 10) для задачи бинарной сегментации артефактов. Для обучения модели использовался предлагаемый набор данных, обучающая и валидационная выборки были разделены в соотношении 3:1. В качестве функции потерь использовалась комбинация бинарной перекрёстной энтропии (BCE) и функция Dice с весами  $\frac{1}{2}$ .

Табл. 1. Сравнение результата работы методов LDL и DeSRA на наборах данных по IoU. Table 1. Comparison of the performance of LDL and DeSRA methods on datasets by IoU.

	DeSRA SR=Real-ESRGAN	Предлагаемый набор данных SR=DRCT
LDL ( $t = 0.01$ )	0.299	0.525
LDL ( $t = 0.005$ )	0.362	0.646
LDL ( $t = 0.001$ )	0.353	0.746
DeSRA ( $t = 0.3$ )	0.511	0.540
DeSRA ( $t = 0.2$ )	0.566	0.622
DeSRA ( $t = 0$ )	–	0.677

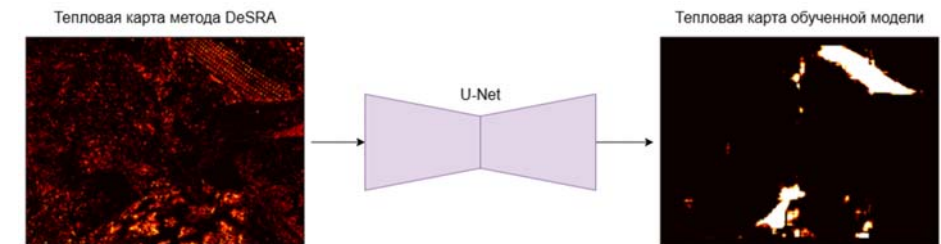


Рис. 10. Схема обученной модели детекции артефактов суперразрешения. Fig. 10. Diagram of the trained model for detecting super-resolution artifacts.

Объективная оценка (табл. 2) модели производилась на наборе данных DeSRA, который имеет три подвыборки для разных методов суперразрешения. Использовалась методология сравнения авторов метода DeSRA. Результаты показывают, что методов суперразрешения Real-ESRGAN [1] и LDL [5] получилось добиться улучшения качества детекции артефактов по показателю IoU. Модель уточняет тепловую карту DeSRA и более явно выделяет области артефактов. Для метода суперразрешения SwinIR [13] обученная модель показывает более плохие результаты, однако это может являться следствием низкого качества разметки набора

данных DeSRA, как было показано выше. На рис. 11 представлены примеры артефактов, которые выделила обученная модель.

Табл. 2. Сравнение предлагаемой модели U-Net с базовым методом DeSRA на наборе данных DeSRA по среднему значению IoU.

Table 2. Comparison of the proposed U-Net model with the baseline DeSRA method on the DeSRA dataset by average IoU.

Подвыборка набора DeSRA по методу суперразрешения	DeSRA	Предлагаемая модель
Real-ESRGAN [1]	0.499	<b>0.517</b>
LDL [5]	0.442	<b>0.478</b>
SwinIR [13]	<b>0.572</b>	0.522
Весь набор DeSRA	0.50639	<b>0.50642</b>



Рис. 11. Примеры артефактов из набора данных DeSRA, детектированных обученной моделью.  
Fig. 11. Examples of artifacts from the DeSRA dataset detected by the trained model.

## 5. Заключение

В работе предложен автоматизированный метод создания наборов данных для детекции артефактов суперразрешения, основанный на применении состязательных атак. Основное улучшение заключалось в модификации функции потерь, которая позволяет создавать локализованные артефакты в заданных областях изображения. Созданный набор данных показал качество разметки выше существующих аналогов. Преимущества метода включают автоматизацию разметки, масштабируемость для разных моделей и возможность контроля

локализации артефактов. В перспективе метод может быть использован для повышения устойчивости моделей суперразрешения к генерации артефактов и созданию качественного метода детекции артефактов.

## Список литературы / References

- [1]. Wang, X., Xie, L., Dong, C., Shan, Y. (2021). Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. International Conference on Computer Vision Workshops (ICCVW).
- [2]. Xie, L., Wang, X., Chen, X., Li, G., Shan, Y., Zhou, J., & Dong, C. (2023). DeSRA: Detect and Delete the Artifacts of GAN-based Real-World Super-Resolution Models.
- [3]. Zhang, L., Zhou, Y., Barnes, C., Amirghodsi, S., Lin, Z., Shechtman, E., Shi, J. (2022). Perceptual artifacts localization for inpainting. European Conference on Computer Vision, 146-164.
- [4]. Zhang, L., Xu, Z., Barnes, C., Zhou, Y., Liu, Q., Zhang, H., Amirghodsi, S., Lin, Z., Shechtman, E., Shi, J. (2023). Perceptual Artifacts Localization for Image Synthesis Tasks. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 7579-7590.
- [5]. Liang, J., Zeng, H., Zhang, L. (2022). Details or Artifacts: A Locally Discriminative Learning Approach to Realistic Image Super-Resolution. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- [6]. Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4), 600-612. DOI: 10.1109/TIP.2003.819861.
- [7]. Ding, K., Ma, K., Wang, S., Simoncelli, E. P. (2022). Image Quality Assessment: Unifying Structure and Texture Similarity. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(5), 2567-2581. DOI: 10.1109/TPAMI.2020.3045810.
- [8]. Zhang, R., Isola, P., Efros, A. A., Shechtman, E., Wang, O. (2018). The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 586-595. DOI: 10.1109/CVPR.2018.00068.
- [9]. Goodfellow, I. J., Shlens, J., Szegedy, C. (2014). Explaining and Harnessing Adversarial Examples. CoRR, abs/1412.6572. Available at: <https://arxiv.org/pdf/1412.6572>, accessed 14.03.2026.
- [10]. Hsu, C.-C., Lee, C.-M., Chou, Y.-S. (2024). DRCT: Saving Image Super-Resolution Away from Information Bottleneck. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 6133-6142.
- [11]. Kuznetsova, A., Rom, H., Alldrin, N. et al. The Open Images Dataset V4. Int J Comput Vis 128, 1956-1981 (2020). DOI: 10.1007/s11263-020-01316-z.
- [12]. Xu, M., Zhang, Z., Wei, F., Hu, H., Bai, X. (2023). SAN: Side Adapter Network for Open-Vocabulary Semantic Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(12), 15546-15561. DOI: 10.1109/TPAMI.2023.3311618.
- [13]. J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool R. Timofte, "SwinIR: Image Restoration Using Swin Transformer", 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 2021, pp. 1833-1844, DOI: 10.1109/ICCVW54120.2021.00210.

## Информация об авторах / Information about authors

Кирилл Владимирович МАЛЫШЕВ получил степень магистра по прикладной математике и информатике в Московском государственном университете имени М. В. Ломоносова в 2023 году. В настоящее время является аспирантом в лаборатории компьютерной графики и мультимедиа МГУ и младшим научным сотрудником в центре искусственного интеллекта МГУ. Также Кирилл работает над стандартом видеокodирования следующего поколения в рамках объединённой команды экспертов по видео (JVET). В область его научных интересов входят методы сжатия видео, оценка качества изображений и видео, нейросетевые методы обработки изображений и видео.

Kirill Vladimirovich MALYSHEV received his M.S. degree in applied mathematics and computer science from Lomonosov Moscow State University in 2023. He is currently a postgraduate student at the Graphics & Media Lab and a junior researcher at the MSU AI Center. Kirill also contributes

to the next-generation video coding standard within the Joint Video Experts Team (JVET). His research interests include video compression methods, image and video quality assessment, and neural network-based image and video processing.

Иван Андреевич МОЛОДЕЦКИХ окончил аспирантуру по специальности математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей в Московском государственном университете имени М. В. Ломоносова в 2024 году. В настоящее время работает над диссертацией и является научным сотрудником в центре искусственного интеллекта МГУ, лаборатории компьютерной графики и мультимедиа. В область его научных интересов входят суперразрешение, семантическое матирование видео и машинное обучение. Иван курировал разработку бенчмарка методов суперразрешения для улучшения качества видео и был одним из организаторов соревнований по оценке качества суперразрешения видео на ECCV-AIM 2024 и ICCV-AIM 2025.

Ivan Andreevich MOLODETSKIKH completed his postgraduate coursework in computer science in the Moscow State University in 2024. He is currently working on his dissertation and is a researcher at the MSU AI Center, Graphics & Media Lab. His research interests include super-resolution, semantic video matting and machine learning. Ivan supervised the development of MSU video upscalers and SR+codec benchmarks and was one of the organizers of the ECCV-AIM 2024 and the ICCV-AIM 2025 video super-resolution quality assessment challenges.

Дмитрий Сергеевич ВАТОЛИН закончил ВМК МГУ в 1996, защитил диссертацию в 2000, кандидат физико-математических наук, заведующий лабораторией компьютерной графики ВМК МГУ. Специализируется на исследованиях в области алгоритмов сжатия видео, современных методах измерения качества и обработке цифрового видео. Читает курсы по компьютерной графике и методам сжатия и обработки видео с 1997 года. Создатель популярных сайтов, посвященных алгоритмам обработки и сжатия видео.

Dmitriy Sergeevich VATOLIN – Cand. Sci. (Phys.-Math.). He received his degree in 2000 from Moscow State University and is currently head of the CS MSU Graphics & Media Lab. His research interests include compression methods, video processing, and 3D-video techniques (depth from motion, focus and other cues, video matting, background restoration, and high-quality stereo generation), as well as 3D-video quality assessment (metrics for 2D-to-3D-conversion artifacts, temporal asynchrony, swapped views, and more). He is the creator of popular websites devoted to video processing and compression algorithms.