



ОТКРЫТАЯ НАУКА
издательство

Международный журнал информационных технологий и энергоэффективности

Сайт журнала:

<http://www.openaccessscience.ru/index.php/ijcse/>



УДК 004.932.4

АПСКЕЙЛ ФОТОМАТЕРИАЛОВ С ПРИМЕНЕНИЕМ НЕЙРОННЫХ СЕТЕЙ

¹Коровин Н.А., ²Коровин М.А., ³Коровин Н.А.

ФГАОУ ВО СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ, Ставрополь, Россия (355017, Ставропольский край, город Ставрополь, ул. Пушкина, д. 1а), e-mail: ¹nikki01946@gmail.com, ²bigsky5068@gmail.com, ³nikki01945@gmail.com

В статье рассматриваются этапы развития технологий апскейлинга: от первых попыток улучшения телевизионного сигнала в 1980-х годах до появления методов суперразрешения (super-resolution) на базе глубокого обучения. Особое внимание уделено истории применения нейронных сетей для улучшения фотоматериалов, включая ключевые прорывы (SRCNN, VDSR, EDSR, ESRGAN и др.) и разнообразие подходов – свёрточные сети, резидуальные сети, генеративно-сопоставительные сети и трансформеры. Описаны принципы работы этих методов. Представлены примеры визуальных результатов до и после апскейлинга, а также схемы работы нейросетевых моделей. Наконец, обсуждаются основные области применения технологий апскейлинга (восстановление старых фотографий и видео, медицине, видеоиграх и др.), а также делаются выводы о значении этих достижений для индустрии обработки изображений.

Ключевые слова: апскейлинг; суперразрешение; повышение разрешения; обработка изображений; нейронные сети; глубокое обучение; свёрточная сеть; масштабирование изображений.

UPSCALE OF PHOTOGRAPHIC MATERIALS USING NEURAL NETWORKS

¹Korovin N.A., ²Korovin M.A., ³Korovin N.A.

NORTH CAUCASUS FEDERAL UNIVERSITY, Stavropol, Russia (355017, Stavropol Territory, Stavropol, Pushkin St., 1a), e-mail: ¹nikki01946@gmail.com, ²bigsky5068@gmail.com, ³nikki01945@gmail.com

This article excuses the stages of development of upscaling technologies: from the first attempts to improve television signals in the 1980s to the advent of super-resolution methods powered by deep learning. Special attention is paid to the history of using neural networks to enhance photographic materials, including key breakthroughs (SRCNN, VDSR, EDSR, ESRGAN, etc.) and the variety of approaches – convolutional networks, residual networks, generative adversarial networks, and transformers. The principles of operation of these methods are described. Examples of visual results before and after upscaling are presented, as well as diagrams illustrating the workings of neural network models. Finally, key application areas of upscaling technology are discussed (restoration of old photos and videos, medicine, video games, etc.), and conclusions are drawn about the significance of these advances for the image processing industry.

Keywords: upscaling; super-resolution; image resolution enhancement; image processing; neural networks; deep learning; convolutional network; image scaling.

Апскейлинг (англ. upscaling) – это процесс увеличения разрешения или масштаба цифрового изображения (фотографии, видеокadra и т.д.) с целью улучшения его четкости и детализации. Иными словами, апскейлинг позволяет преобразовать исходный низкокачественный визуальный материал в более качественный за счет добавления недостающей информации. Современные алгоритмы способны достраивать изображение – добавлять отсутствующие пиксели таким образом, чтобы итоговое изображение выглядело максимально натурально и четко. Данная технология сегодня находит широкое применение:

от улучшения старых фильмов и фотографий до масштабирования графики видеоигр и повышения качества видео с камер наблюдения. Развитие методов апскейлинга стало неотъемлемой частью цифровой эпохи, позволяя сохранять высокий уровень детализации даже при работе с изначально низким разрешением контента [1]. В следующих разделах рассмотрим, как возник и развивался апскейлинг, и какую роль в его прогрессе сыграли нейронные сети.

Ранние методы апскейла появились в 1980–2000-е гг. Концепция повышения разрешения изображений зародилась еще в 1980-х годах как теория обработки сигналов [5]. Первые практические реализации апскейлинга были связаны с телевидением: для улучшения качества аналогового ТВ-сигнала применяли простейшие алгоритмы масштабирования, например линейную интерполяцию, что давало лишь незначительное повышение четкости без серьезных искажений. Также разрабатывались методы комбинирования нескольких кадров одной сцены (множественная съемка) для извлечения дополнительных деталей – так называемое мультифреймовое суперразрешение, требовавшее сложной оценки движения между кадрами. Эти классические подходы были вычислительно затратны и чувствительны к шуму и неточному выравниванию изображений.

С развитием цифровых технологий в 1990-е годы апскейлинг значительно продвинулся. Появились более сложные математические модели обработки: от алгоритмов фильтрации и реставрации изображения до методов на основе статистических предположений (регуляризация) и фрактального увеличения разрешения. Произошёл переход от аналоговых средств улучшения картинки к цифровым, что позволило внедрять апскейлинг в DVD-плееры и телесигналы высокого разрешения. Например, начали применяться бикубическая интерполяция и другие усовершенствованные алгоритмы ресемплинга, которые учитывали больше соседних точек при расчете новых пикселей по сравнению с простым линейным масштабированием. Тем не менее все эти методы оставались ограничены тем, что они основывались на фиксированных правилах обработки пиксельных данных и не могли придумать новые детали, отсутствующие в оригинале [1].

Прорывом в области повышения качества изображений стали примерно-образные методы (example-based super-resolution). В 2003 году Уильям Фримен с коллегами предложили алгоритм, использующий обучающую выборку пар изображений низкого и высокого разрешения для улучшения качества. Их метод искал в базе фрагменты (патчи) изображений, сходные с участками увеличиваемого фото, и заменял их более детализованными фрагментами высокого разрешения. Этот подход на основе обученных примеров стал первым шагом к использованию обучаемых моделей в задаче апскейлинга вместо жестко запрограммированных правил. Появились алгоритмы, полагающиеся на статистику больших наборов изображений, что позволило лучше восстанавливать текстуры и мелкие детали по сравнению с одной лишь интерполяцией. Тем не менее, вычислительно такие методы были сложны, а качество результата сильно зависело от репрезентативности базы образов.

Начало 2010-х ознаменовало новый этап: в обработке изображений получили развитие алгоритмы глубокого обучения, в частности свёрточные нейронные сети (Convolutional Neural Networks, CNN). Впечатляющий прогресс в смежных задачах компьютерного зрения натолкнул исследователей на мысль применить нейросети и для апскейлинга. Ключевая идея заключалась в том, чтобы обучить сеть восстанавливать изображение высокого разрешения по его низкоразрешенной версии, используя большие объемы данных. В 2014 году была

представлена первая модель глубокого обучения для суперразрешения – SRCNN (Super-Resolution CNN), предложенная Чао Дунгом и его коллегами. Несмотря на относительную простоту (всего три слоя), SRCNN смогла повысить качество увеличения изображений до уровня, значительно превосходящего традиционные методы интерполяции. Фактически, сеть научилась восстанавливать утраченные детали, опираясь на знание, приобретённое при обучении на большом наборе примеров. Успех SRCNN породил всплеск интереса к нейросетевому апскейлингу и стимулировал появление новых архитектур, ещё более совершенствующих результат [5].

Уже через пару лет глубина и сложность моделей выросли. В 2016 году была предложена очень глубокая свёрточная сеть VDSR (Very Deep Super Resolution) с 20 слоями, сумевшая достичь ещё большего прироста качества за счёт увеличения сети и введения резидуального обучения. В VDSR исходное изображение складывается с выходным сигналом сети, что позволяет ей учить только разницу (детали высокого частотного диапазона) между высоким и низким разрешением. Такой подход заметно упростил обучение глубоких сетей, устранив проблему затухающих градиентов, и улучшил точность восстановления мелких структур [6]. Далее последовали улучшенные архитектуры: в 2017 году – EDSR (Enhanced Deep SR), упростившая некоторые слои и убравшая ограничения нормализации для увеличения качества, а также модели с механизмами внимания. Параллельно были предприняты попытки сделать картинки визуально более реалистичными: вместо оптимизации только математических метрик ошибки исследователи начали применять генеративно-состязательные сети (GAN).

Одной из первых таких моделей стал SRGAN (2017 г.), в которой сеть-генератор училась создавать фотореалистичные детали, пытаясь обмануть сеть-дискриминатор, обученную отличать увеличенные изображения от настоящих высококачественных фотографий. Развивая эту идею, в 2018 году представлена модель ESRGAN (Enhanced SRGAN), которая еще лучше справлялась с восстановлением текстур – например, деталей кожи или травы – делая увеличенные изображения практически неотличимыми на глаз от оригинальных по резкости и нюансам. Одновременно продолжалось совершенствование прямых методов: увеличивалось количество слоев, вводились механизмы self-attention и другие улучшения архитектур.

В начале 2020-х годов в задаче суперразрешения стали применять и трансформерные нейросети – изначально разработанные для задач NLP модели внимания оказались эффективны и для обработки изображений. С их помощью удаётся учитывать дальние связи в изображении и еще лучше вписывать мелкие детали в общий контекст (примеры – модели типа IPT, SwinIR и др., применяемые в отдельных доменах, например для медицинских снимков) [5].

Получается, что за последнее десятилетие методы апскейлинга с помощью искусственного интеллекта прошли путь от первых экспериментальных сетей к целому семейству мощных алгоритмов, способных выдавать потрясающие результаты. Ниже рассмотрим более подробно, какие виды нейросетевых подходов к повышению разрешения существуют и как они работают.

Нейросетевой апскейлинг имеет различия по типам и принципам.

Single Image vs. Multi-Image – большинство современных нейросетевых методов ориентированы на повышение разрешения одного изображения – так называемое Single Image Super-Resolution (SISR). В этой задаче на вход подаётся одиночное фото низкого разрешения,

и сеть генерирует его высокоразрешенную версию. Однако есть и многокадровые варианты: видео-суперразрешение, где сеть обрабатывает последовательность кадров, учитывая информацию о движении, чтобы повысить чёткость видео без артефактов на стыках кадров. Например, такие алгоритмы используются при оцифровке и ремастеринге старых фильмов – каждый кадр улучшается, сохраняя плавность переходов между ними. Кроме того, существуют мультиспектральные и мультимодальные подходы, когда помимо обычного изображения нейросеть принимает дополнительную информацию (например, инфракрасный канал или карту глубины) для повышения качества – это позволяет добиваться лучшего результата за счёт комбинации разных данных, хотя такие методы применяются в узких областях (например, в спутниковой съемке).

Архитектуры на основе CNN – базой большинства решений для апскейлинга служат свёрточные нейронные сети. Простая архитектура типичной модели суперразрешения состоит из последовательности свёрточных слоёв, которые постепенно преобразуют размытое низкокачественное изображение в чёткое высококачественное. Рисунок 1 иллюстрирует принцип работы такой сети на примере оригинального алгоритма SRCNN.

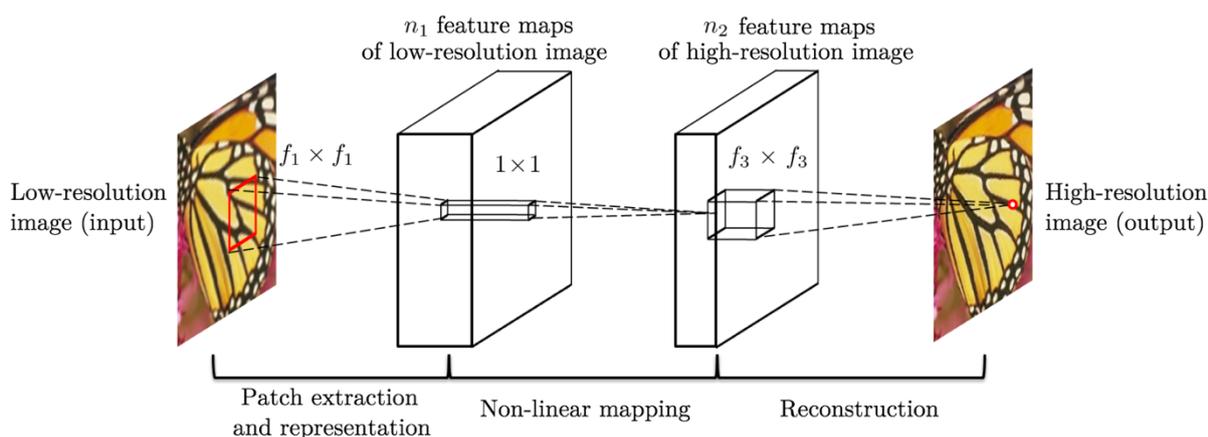


Рисунок 1 – Упрощенная схема работы свёрточной нейросети SRCNN для повышения разрешения: входное низкого качества изображение пропускается через несколько слоёв (патч-экстракцию признаков, нелинейное преобразование и реконструкцию), в результате чего на выходе формируется изображение повышенного разрешения. Такая сеть обучается по большим наборам пар изображений (LR→HR) и способна восстанавливать утраченные детали.

Первый свёрточный слой в подобных сетях выделяет из размытого изображения базовые черты (границы, текстуры) – фактически извлекает множество признаков низкого уровня. Последующие слои постепенно строят более сложные характеристики, сопоставляя полученные признаки с шаблонами, усвоенными из обучающей выборки. Наконец, финальный слой генерирует итоговое изображение более высокого разрешения, объединяя уточненные детали с общей структурой изображения. В примере SRCNN использовалось всего 3 слоя (9×9 , 1×1 и 5×5 фильтры), однако даже такая небольшая сеть показала заметное улучшение по сравнению с интерполяцией. Более глубокие сети, например упомянутая VDSR из 20+ слоёв, идут дальше: они способны восстанавливать очень тонкие детали (например, текст или поры кожи на фото), но для их успешного обучения требовались дополнительные

приемы. Одно из ключевых усовершенствований – резидуальные (остаточные) связи. Добавление резидуальной связи означало, что сеть учит не всё изображение целиком, а лишь разницу между высоким и низким разрешением. На практике это позволило сосредоточить ресурсы на восстановлении мелких деталей, избегая повторного обучения тому, что уже присутствует во входном изображении (большие однородные области, низкочастотные компоненты и т.д.). Резидуальные архитектуры (ResNet) не только повысили качество, но и упростили обучение очень глубоких сетей, сделав возможным использование десятков и даже сотен слоёв без деградации градиента [6].

Генеративные модели и GAN. Одно из ограничений классических CNN-методов суперразрешения – стремление максимизировать показатели вроде PSNR (Peak Signal-to-Noise Ratio), что зачастую приводит к сглаживанию изображения. Сети, оптимизируемые по MSE (среднеквадратичной ошибке), восстанавливают «усреднённые» детали и могут давать слишком мягкую картинку. Чтобы получить более фотореалистичный результат, исследователи стали применять генеративные подходы. Наибольшую популярность получили генеративно-сопоставительные сети (GAN), состоящие из двух моделей: генератора, который пытается создать правдоподобное высококачественное изображение, и дискриминатора, задача которого – отличить сгенерированное изображение от настоящего высокого разрешения. В процессе совместного обучения генератор постепенно учится «обманывать» дискриминатор, добавляя всё более реалистичные высокочастотные детали. Применительно к апскейлингу такой подход впервые реализован в модели SRGAN (Ledig и др., 2017), а улучшен – в ESRGAN (Wang и др., 2018). Модель ESRGAN смогла генерировать текстуры (например, траву, волосы, кирпичную кладку) с высокой достоверностью, сильно приближая визуальное качество увеличенного изображения к оригиналу [5]. Однако GAN-методы имеют и недостатки: порой они могут «домысливать» детали слишком агрессивно, что приводит к появлению мелких артефактов или чересчур резких элементов, отсутствующих в оригинале. В практических приложениях часто комбинируют подходы – например, используют гибридные функции потерь: смесь перцептуальных критериев и традиционной MSE, чтобы добиться баланса между четкостью и достоверностью.

Ключевым фактором успеха глубинного апскейлинга стала доступность больших данных и возросшие вычислительные мощности. Нейросеть «учат» на паре изображений: исходном высококачественном и его искусственно уменьшенной копии (именно такую пару называют Teacher и Student изображениями). Огромные наборы таких пар (тысячи и миллионы изображений самых разных сцен) используются для тренировок – сеть постепенно подбирает параметры фильтров, минимизируя разницу между своим выходом и эталонным высоким разрешением. В процессе обучения нейронная сеть фактически запоминает паттерны реального мира: формы объектов, текстуры поверхностей, характерные шумы камеры и т.д. Например, видя тысячи портретов, сеть научается достраивать резкие глаза и ресницы даже на размытой фотографии, а анализируя пейзажи – узнаёт, как выглядит листва или трава при большом увеличении. В результате современный алгоритм апскейлинга – это не просто математическая интерполяция, а сложная модель, обладающая обобщёнными знаниями о разных типах изображений. Тем не менее, важно понимать ограничения: нейросеть не «угадывает» реальные отсутствующие детали, а генерирует правдоподобные детали по аналогии с тем, что видела на обучающих данных. Поэтому иногда возникают ситуации, когда пользователи ошибочно принимают сгенерированную детализацию за настоящую – например,

улучшенное ИИ фото может казаться раскрывающим черты лица, которых нет на исходном размытом снимке [3]. В критичных приложениях (например, обработка снимков в криминалистике) этот фактор обязательно учитывается.



Рисунок 2 – Фотография Трампа, усиленная искусственным интеллектом, которая стала вирусной на фоне слухов о плохом здоровье

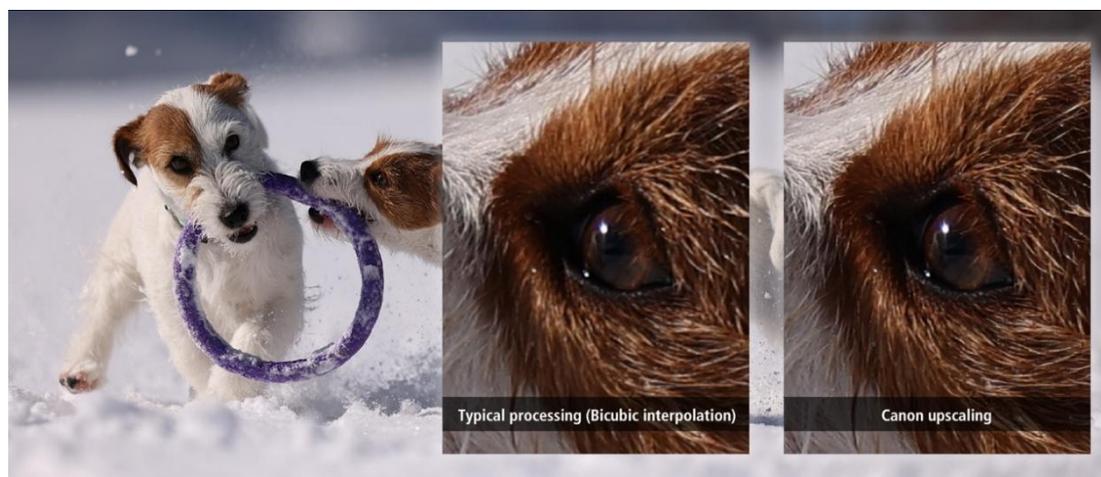


Рисунок 3 – Сравнение результатов традиционного бикубического увеличения (слева) и апскейлинга с помощью нейронной сети (справа) на фрагменте фотографии собаки. Видно, что нейросетевой метод сохраняет тонкую детализацию шерсти и границ лучше, тогда как при обычном масштабировании изображение размыто и теряет резкость.

Для конечного пользователя современные инструменты апскейлинга часто представляют собой «чёрный ящик», выполняющий сложные вычисления по нажатию кнопки. Благодаря оптимизациям и использованию аппаратного ускорения (GPU, TPU), нейросетевые модели способны работать практически в реальном времени. К примеру, в игровых приложениях технология DLSS (Deep Learning Super Sampling) от NVIDIA генерирует кадры

высокой чёткости из низкого разрешения прямо на лету, позволяя значительно повысить производительность без заметной потери качества картинки [4]. Аналогично, смартфоны при съемке могут автоматически улучшать детали фото с помощью встроенных моделей ИИ. Таким образом, нейросетевой апскейлинг перестал быть сугубо лабораторной разработкой – он внедряется в массовые продукты, зачастую незаметно для самого пользователя.

Сегодня методы увеличения разрешения с помощью ИИ широко применяются в самых разных областях. В медицине они помогают улучшать качество диагностических изображений (например, МРТ, КТ) для более точного анализа. В фотографии и кинематографе нейросети используются для реставрации и ремастеринга архивных материалов – старые фильмы, снятые в низком разрешении, обретают новую жизнь в формате HD и 4K, открывая зрителю ранее невидимые детали. Энтузиастами и профессионалами активно восстанавливаются старые фотоснимки: алгоритмы удаляют шум, повышают четкость лиц на исторических фотографиях, что было невозможно традиционными методами. В видеонаблюдении суперразрешение помогает разобрать номера автомобилей или лица в кадрах с камер низкого качества – это повышает эффективность систем безопасности. Отдельно стоит отметить сферу развлечений: помимо уже упомянутых видеоигр, апскейлинг используется в приложениях дополненной и виртуальной реальности для оптимизации рендеринга.

Важно подчеркнуть, что качество работы разных методов апскейлинга может различаться в зависимости от типа изображения. Например, для анимационной графики разработаны специализированные алгоритмы (такие как Waifu2x для аниме-артов [7]), учитывающие особенности рисунка и плоских областей цвета. Для лиц человека существуют модели, обученные на портретах, способные достоверно восстанавливать глаза, рот и прочие черты. В общем случае универсальные алгоритмы стараются быть максимально адаптивными. Современные тенденции в исследовании суперразрешения включают объединение подходов (например, каскадное применение нескольких моделей, где одна устраняет шум, другая увеличивает размер), а также изучение методов обучения без учителя для случаев, когда нет идеальных пар изображений для тренировки.

Прогресс в области нейронного апскейлинга не только улучшил метрики качества изображений, но и снизил порог входа для использования этих технологий. Если раньше для подобных задач требовалось специализированное оборудование и месяцы обучения модели, то теперь доступны готовые облачные сервисы, позволяющие вызвать суперразрешение через API или встроить в программный продукт [5]. Это открыло дорогу множеству приложений: от улучшения пользовательских фотографий в социальных сетях до масштабных проектов оцифровки библиотек изображений и видеоархивов. Апскейлинг с помощью ИИ превратился из научного эксперимента в повседневный инструмент, значительно расширяющий возможности работы с визуальной информацией.

История развития технологий повышения разрешения наглядно демонстрирует, как искусственный интеллект открыл новую сферу обработки изображений. Нейросетевые алгоритмы апскейлинга за короткий срок преодолели ограничения классических подходов, научившись восстанавливать детали там, где раньше воображение упиралось в размытые пиксели. Сегодня суперразрешение служит связующим звеном между растущими запросами на качество визуального контента и ограничениями исходных данных – будь то старый фотоматериал или ресурсоограниченные устройства. Достижения в этой области не только дают эстетическое улучшение изображений, но и несут практическую ценность: делают

возможным анализ ранее непригодных материалов, улучшают точность медицинской диагностики, сохраняют культурное наследие в цифровом виде.

Список литературы

1. Умнова Александра. Апскейл – что такое. // Skyeng [Электронный ресурс] URL: <https://skyeng.ru/magazine/wiki/it-industriya/chto-takoe-apskeil/> (Дата обращения: 12.11.25 г.)
2. Технология масштабирования нейронных сетей // Canon Global Tech Article, 13.11.2023 [Электронный ресурс] URL: <https://global.canon/en/technology/dl-upscaling-2023.html> (Дата обращения: 12.11.25 г.)
3. Гроукут Мэтт. Фотография Трампа демонстрирует опасности, связанные с расширением возможностей искусственного интеллекта. // PetaPixel. [Электронный ресурс] URL: <https://petapixel.com/2025/09/05/trump-photo-shows-the-perils-of-ai-upscaling-tools/> (Дата обращения: 12.11.25 г.)
4. NVIDIA DLSS: Ваши вопросы, ответы. // Официальный сайт Nvidia [Электронный ресурс] URL: <https://www.nvidia.com/en-us/geforce/news/nvidia-dlss-your-questions-answered/> (Дата обращения: 12.11.25 г.)
5. Рост сверхразрешения: от классической обработки изображений до масштабируемых API искусственного интеллекта // IBM [Электронный ресурс] URL: <https://community.ibm.com/community/user/blogs/paul-glenn2/2025/04/25/the-rise-of-super-resolution-from-classical-image#:~:text=The%20concept%20of%20super,sensitive%20to%20noise%20and%20misalignment> (Дата обращения: 12.11.25 г.)
6. Сверхрезультативность // Bayern Collab [Электронный ресурс] URL: <https://collab.dvb.bayern/spaces/TUMlfdv/pages/69119923/Super-Resolution> (Дата обращения: 12.11.25 г.)
7. Еще более качественное масштабирование изображений с помощью Waifu2x. // Журнал Fedora Magazine. 2018-10-02. [Электронный ресурс] URL: <https://fedoramagazine.org/better-image-upscaling-waifu2x/> (Дата обращения: 12.11.25 г.)

References

1. Umnova Alexandra. Upscale – what is it. // Skyeng [Electronic resource] URL: <https://skyeng.ru/magazine/wiki/it-industriya/chto-takoe-apskeil/> (Date of request: 12.11.25)
2. Neural network scaling technology // Canon Global Tech Article, 11/13/2023 [Electronic resource] URL: <https://global.canon/en/technology/dl-upscaling-2023.html> (Date of request: 12.11.25)
3. Growcoot Matt. Trump Photo Shows The Perils of AI Upscaling Tools. // PetaPixel. [Электронный ресурс] URL: <https://petapixel.com/2025/09/05/trump-photo-shows-the-perils-of-ai-upscaling-tools/> (Дата обращения: 12.11.25 г.)
4. NVIDIA DLSS: Your Questions, Answered. // Official Nvidia website [Электронный ресурс] URL: <https://www.nvidia.com/en-us/geforce/news/nvidia-dlss-your-questions-answered/> (Дата обращения: 12.11.25 г.)
5. The Rise of Super-Resolution: From Classical Image Processing to Scalable AI APIs // IBM [Электронный ресурс] URL: <https://community.ibm.com/community/user/blogs/paul->

