

Научная статья

УДК 621.397

<https://doi.org/10.31854/1813-324X-2025-11-2-7-19>

EDN:TKAPTM



Гибридный метод локального контрастирования изображений с нейросетевой регулировкой параметров

✉ Иван Юрьевич Грицкевич, gritskevich.iu@sut.ru

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

Аннотация

Актуальность. Современные методы обработки изображений направлены на повышение их визуального качества, в частности, на адаптивное локальное контрастирование. Для достижения высокой эффективности контрастирования ранее применялись классические алгоритмы, однако они не учитывали глобальный контекст сцены и могли приводить к усилению шумовых искажений. В связи с этим в данной работе предложен гибридный метод адаптивного локального контрастирования изображений с использованием нейросетевой регулировки параметров.

Целью статьи является разработка алгоритма, обеспечивающего оптимальное усиление контраста при минимизации шумовых артефактов и искажений, повышение контрастности и точности обнаружения объектов в режиме реального времени.

Сущность решения: адаптивная настройка начальных параметров локального контрастирования с помощью сверточной нейронной сети, учитывающей яркостные и текстурные особенности. Сверточная нейронная сеть динамически подбирает параметры обработки и размеры локальных областей для объектов и фона, улучшая видимость деталей и подавляя артефакты обработки (ореолы, блочность). Метод реализован в виде программно-аппаратного комплекса для компьютерного зрения, обработки аэрофотоснимков, видеонаблюдения и поиска пострадавших при катаклизмах.

Научная новизна работы заключается в разработке алгоритма, позволяющего автоматически регулировать параметры контрастирования на основе анализа глобального и локального контекста сцены с использованием искусственного интеллекта.

Теоретическая значимость работы состоит в гибридном подходе к адаптивной обработке изображений, основанном на применении сверточной нейронной сети для управления параметрами локального контрастирования. Управление параметрами осуществляется на основе анализа текстурных и частотных характеристик изображения, автоматическую адаптацию под которые производит нейронная сеть. Методика обеспечивает адаптацию к нестационарным условиям наблюдения и, как следствие, повышает устойчивость алгоритма к сложным условиям.

Практическая значимость разработанного алгоритма определяется реализацией повышения контраста объектов изображений, полученных в видимом и инфракрасном диапазонах спектра и достоверностью их распознавания с использованием искусственного интеллекта.

Ключевые слова: контраст, искажения, оценка качества изображений, обработка изображений, нейронные сети, частотные характеристики, локальная оценка, адаптация, локальные особенности

Ссылка для цитирования: Грицкевич И.Ю. Гибридный метод локального контрастирования изображений с нейросетевой регулировкой параметров // Труды учебных заведений связи. 2025. Т. 11. № 2. С. 7–19. DOI:10.31854/1813-324X-2025-11-2-7-19. EDN:TKAPTM

Original research

<https://doi.org/10.31854/1813-324X-2025-11-2-7-19>

EDN:ТКАРТМ

A Hybrid Approach to Local Contrast Enhancement Using Adaptive Neural Network Parameter Control

Ivan Yu. Gritskevich, gritskevich.iu@sut.ruThe Bonch-Bruевич Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Annotation

Relevance. Modern image processing techniques are focused on enhancing visual quality, particularly through adaptive local contrast enhancement. Previously, classical algorithms were employed to achieve high contrast efficiency; however, these approaches failed to account for the global scene context and often led to noise amplification. This paper proposes a hybrid method for adaptive local image contrast enhancement utilizing neural network-based parameter adjustment.

The **aim** of this research is to develop an algorithm that provides optimal contrast enhancement while minimizing noise artifacts and distortions, thereby improving contrast and real-time object detection accuracy.

The **essence** of the proposed solution lies in employing a convolutional neural network for automatic configuration of local contrast parameters based on statistical brightness characteristics and textural image features. The proposed method incorporates image segmentation into local regions, analysis of their properties, and adaptive adjustment of processing parameters. This results in improved discernibility of low-contrast objects under various imaging conditions. The algorithm's operating principle is **based** on dynamically selecting local region dimensions and contrast parameters depending on background and target scene objects. The integration of a neural network module enables precise adjustment of processing parameters while minimizing undesirable artifacts such as halos and blockiness. The methodology has been **implemented** as software and hardware for an optoelectronic system designed for computer vision applications, aerial image processing, video surveillance systems, and locating victims in various disaster scenarios.

The **scientific novelty** of this work lies in the development of an algorithm that automatically regulates contrast parameters based on analysis of both global and local scene context using artificial intelligence.

The **theoretical significance** of the work consists in the development of a contrast enhancement algorithm and image quality assessment method that accounts for contrast perception characteristics by both humans and AI systems under challenging observational conditions, such as fog, smoke, low illumination, etc.

The **practical significance** of the developed algorithm is determined by its implementation of contrast enhancement for objects in images acquired in both visible and infrared spectral ranges, and by the reliability of their recognition using artificial intelligence.

Keywords: contrast, distortion, image quality assessment, image processing, neural networks, frequency characteristics, local evaluation, adaptation, local features

For citation: Gritskevich I.Yu. A Hybrid Approach to Local Contrast Enhancement Using Adaptive Neural Network Parameter Control. *Proceedings of Telecommunication Universities*. 2025;11(2):7–19. (in Russ.) DOI:10.31854/1813-324X-2025-11-2-7-19. EDN:ТКАРТМ

Введение

Современные системы компьютерного зрения, медицинской диагностики, аэрофотосъемки и автономного транспорта требуют качественной обработки изображений для повышения разборчивости объектов и точности алгоритмов анализа. Однако изображения, полученные в сложных усло-

виях, часто страдают от низкой контрастности. Это может быть вызвано различными факторами, такими как неравномерное освещение сцены, когда тени или блики скрывают важные детали [1], ограниченный динамический диапазон сенсора, не позволяющий одновременно захватить яркие и темные области [2], или же свойствами самих материалов, например, низкой отражательной способностью [3].

Низкий контраст приводит к тому, что важные детали становятся неразличимыми, затрудняя обнаружение объектов [4–6], сегментацию изображений [7–9] и 3D-реконструкцию [10–12, 13]. В результате, два близко расположенных объекта могут быть восприняты как один [14].

Эффективное улучшение контраста имеет первостепенное значение для широкого спектра приложений, включая криминалистику, мониторинг дорожного движения, инспекцию промышленной безопасности [15], медицинскую визуализацию для диагностики [3], анализ спутниковых снимков для мониторинга окружающей среды и систем наблюдения, работающие в сложных условиях. В связи с этим задача адаптивного контрастирования становится критически важной для улучшения восприятия и анализа визуальной информации.

Различные методы повышения контрастности уже широко применяются в обработке изображений, однако их эффективность ограничивается рядом недостатков. Хотя глобальные методы просты в реализации [14], они часто не могут устранять локальные вариации контраста внутри изображения.

Применение однородного преобразования ко всему изображению может привести к усилению шума в одних областях и подавлению важных деталей в других [15, 16]. Это связано с тем, что локальное контрастирование, по сути, является полосовым фильтром изображения [17]. Глобальные методы, такие как гистограммная эквализация [7], не учитывают эти локальные частотные характеристики, что и приводит к нежелательным эффектам, к потере деталей и усилению шумов [18].

Локальные методы, включая CLaNE (аббр. от *англ.* Contrast Limited Adaptive Histogram Equalization) [19] и Retinex [8], адаптируются к локальному контексту, но могут вызывать артефакты, такие как блочность и ореолы [1, 3]. К распространенным недостаткам существующих алгоритмов локального контрастирования относятся усиление шума, особенно в однородных областях, появление ореолов вокруг резких краев, искажения цвета и потенциальная потеря мелких деталей, особенно в очень ярких или очень темных областях изображения [15].

Глубокие нейросетевые модели обучаются напрямую преобразовывать пиксели входного изображения в пиксели выходного, минимизируя ошибку по сравнению с эталонным изображением, и при этом сами «выясняют», как лучше всего выполнить это преобразование [12, 20, 21], но они требуют значительных вычислительных ресурсов [2]. Основной проблемой улучшения контраста на основе сверточных нейронных сетей (CNN, аббр. от *англ.* Convolutional Neural Network) является их сильная зависимость от больших объемов изображений с низким контрастом и соответствующих изображений с высоким контрастом [15].

В инфракрасном (ИК) диапазоне значительный вклад в шум вносит фоновое излучение. Это компромиссе между пространственным и температурным (энергетическим) разрешением. Выбор спектрального диапазона влияет на температурный коэффициент чувствительности, поскольку интенсивность излучения черного тела и контраст между объектами и фоном зависят от длины волны.

Фундаментальной проблемой улучшения контраста является достижение значительного увеличения качества изображения без одновременного внесения нежелательных артефактов или потери ценных деталей изображения. Эти ограничения создают необходимость в разработке гибридного метода, объединяющего преимущества традиционных и нейросетевых алгоритмов. Интеграция традиционных методов обработки изображений с современными возможностями обучения моделей искусственного интеллекта предлагает перспективный путь к разработке более адаптивных, и, в конечном итоге – более эффективных методов улучшения изображений [9, 22].

Искусственный интеллект, особенно модели глубокого обучения, такие как CNN, может использоваться для анализа локальных характеристик изображения и последующей адаптивной настройки параметров в традиционных алгоритмах улучшения контраста [21, 23]. CNN фактически выполняет многомасштабный анализ, извлекая признаки на разных уровнях абстракции. Это обеспечивает более контекстно-зависимое применение этих алгоритмов. И наоборот, традиционные методы, которые часто основаны на хорошо изученных математических принципах [17, 24, 25], могут обеспечить уровень надежности и интерпретируемости, которого может не хватать чисто сквозным подходам, управляемым искусственным интеллектом.

Сравнение традиционных методов и CNN приведено в таблице 1, где используются следующие сокращения:

– MSR (аббр. от *англ.* Multi-Scale Retinex) – метод улучшения изображений (особенно контраста и цветопередачи), основанный на теории Retinex и применяющий ее на нескольких пространственных масштабах (с разными размерами окон);

– MSRCR (аббр. от *англ.* Multi-Scale Retinex with Color Restoration) – расширение MSR, добавляющее шаг восстановления цвета для коррекции искажений, которые может вносить базовый MSR;

– SSR (аббр. от *англ.* Single-Scale Retinex) – базовый вариант Retinex, применяемый только на одном пространственном масштабе (с одним размером окна) для улучшения контраста;

– HE (аббр. от *англ.* Histogram Equalization) – метод глобального контрастирования изображения путем выравнивания его гистограммы распределения яркостей.

ТАБЛИЦА 1. Сравнение методов контрастирования

TABLE 1. Comparison of Contrast Methods

Метод	Принцип	Преимущества	Ограничения	Типичные приложения	Вычислительная сложность
Контрастно-ограниченное адаптивное выравнивание гистограммы (CLAHE)	Локальное выравнивание гистограммы с ограничением контраста	Улучшение локального контраста, ограничение усиления шума, уменьшение артефактов на границах	Более сложная реализация по сравнению с HE, требует настройки параметров (размер плитки, предел отсечения)	Медицинская визуализация, подводная фотография, низкая видимость	Средняя
Многомасштабный Retinex (MSR)	Взвешенная сумма результатов SSR в разных масштабах	Лучший баланс между сжатием динамического диапазона и цветопередачей по сравнению с SSR	Может приводить к цветовым искажениям, требует выбора нескольких масштабов	Улучшение контраста и цветопередачи в различных условиях освещения	Высокая
Многомасштабный Retinex с восстановлением цвета (MSRCR)	MSR с дополнительными шагами восстановления цвета	Улучшение контраста и цветопередачи с уменьшением цветовых искажений	Может инвертировать цвета в насыщенных изображениях, требует настройки дополнительных параметров	Улучшение контраста и цветопередачи изображений, не соответствующих предположению о «сером мире»	Высокая
Сверточные нейронные сети (CNN)	Обучение сложным отображениям между изображениями с низким и высоким контрастом	Высокая эффективность, способность к изучению сложных признаков, адаптивность к различным задачам	Требуют больших объемов данных для обучения, значительных вычислительных ресурсов, интерпретируемость может быть затруднена	Улучшение изображений с низкой освещенностью (подводных, медицинских)	Очень высокая (обучение), высокая (вывод)

Одним из ключевых принципов современной обработки изображений является многомасштабный анализ, позволяющий учитывать как локальные особенности, так и глобальный контекст сцены. Этот принцип находит отражение в различных подходах, включая фрактальный анализ, вейвлет-преобразования и нейросетевые модели, извлекающие признаки на разных уровнях абстракции [6, 25]. Объединение этих сильных сторон классических подходов и современных нейросетевых алгоритмов приводит к более надежным, адаптивным и потенциально интерпретируемым решениям для локальной регуляции контраста.

Чтобы увеличить контраст гибридным методом, требуется решение следующих задач:

- разработка алгоритма с адаптивным ограничением контраста и динамическим выбором размера окна обработки;
- применение CNN для автоматической настройки параметров контрастирования;
- интеграция хроматического контрастного веса, минимизирующего цветовые искажения;
- оптимизация метода для реализации в системах реального времени (включая программируемую логическую интегральную схему (ПЛИС));
- экспериментальное тестирование на различных наборах данных (ИК изображения, аэрофото-съемка и т. д.);
- сравнение с существующими методами.

Особое внимание в статье уделено обработке ИК изображений, поскольку в этом спектральном диапазоне контрастность сцены зависит от темпера-

турных характеристик объектов. Это приводит к дополнительным проблемам, таким как температурная зависимость, фиксированные шумы сенсоров и атмосферные искажения [10, 11]. Применение в ИК съемке наиболее важно для задач поиска и спасения людей, поскольку позволяет выделить объекты с разной температурой. В ИК диапазоне контраст зависит от теплового излучения, что влияет на восприятие объектов, однако стандартные методы, например CLAHE, усиливают шум, что делает анализ менее достоверным. Помимо внешних факторов (шум сенсора и атмосферные искажения) на качество восприятия изображений влияет внутренний шум зрительной системы [6]. Этот шум обусловлен различными процессами, включая спонтанную активность и тепловые флуктуации.

Разработка адаптивного гибридного метода позволила объединить преимущества локального контрастирования и нейросетевых технологий, обеспечивая устойчивую работу в различных условиях без ручной настройки параметров, минимизацию артефактов (блочности, ореолов, усиления шума) и оптимальную вычислительную эффективность для работы в реальном времени. Таким образом, разработанный алгоритм предоставляет перспективное решение для повышения качества изображений в различных областях применения.

Предложенный метод включает несколько последовательных этапов.

Этап 1. Выбор начальных параметров скользящего окна, исходя из обстановки либо исходя из предыдущего кадра.

Эман 2. Анализ статистических характеристик внутри скользящего окна и его изменение во время обработки.

Эман 3. Адаптивное усиление контраста.

Эман 4. Оценка всего кадра с помощью собственной безэталонной метрики для регулировки последующего кадра.

Эман 5. Регулировка параметров обработки с помощью нейросетевой модели.

Разработанный алгоритм локального контрастирования имеет несколько настраиваемых параметров:

- размер окна (динамический, ограниченный диапазоном $[r_{\min}, r_{\max}]$);
- порог отсеечения гистограммы T_c ;
- коэффициенты в критериях, используемых для адаптации размера окна и порога отсеечения.

В ходе исследования проведен анализ, как выбор этих параметров влияет на качество выходного изображения.

Динамическое скользящее окно

При максимальном размере окна (рисунок 1) ($r_{\max} \rightarrow$ размер изображения) локальное контрастирование вырождается в глобальное. Таким образом, выходная яркость $I'(x, y)$ всех пикселей определяется глобальной функцией преобразования, рассчитанной по гистограмме всего изображения. Это максимизирует общий контраст, но теряет локальные детали. При минимальном окне ($r_{\min} = 0$) локальная гистограмма вырождается, и $I'(x, y) \approx I(x, y)$.

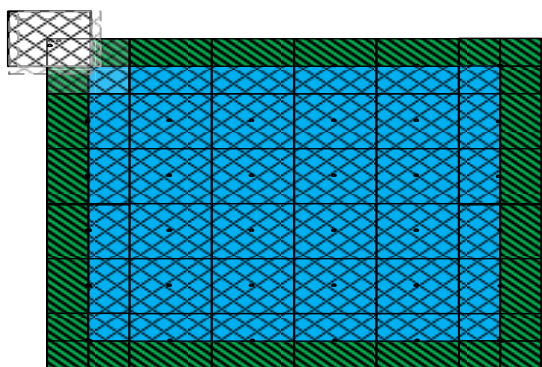


Рис. 1. Скользящее окно по кадру

Fig. 1. Sliding Window for Local Neighborhood Analysis

Если размер окна $r(x, y)$ кусочно-постоянен, то внутри области с $r(x, y) = r_0$ локальное преобразование $T_{\Omega r_0}$ эквивалентно свертке с пространственно-инвариантным оператором. $I'(x, y)$ зависит только от пикселей в окрестности $\Omega r_0(x, y)$ и не зависит от абсолютного положения (x, y) внутри области. Это преобразование непрерывно и монотонно, не создавая разрывов градиента.

На границе областей с разными размерами окон r_1 и r_2 соседние пиксели (x, y) и $(x + 1, y)$ обрабаты-

ваются в окрестностях $\Omega r_1(x, y)$ и $\Omega r_2(x + 1, y)$. Разность их выходных значений $|I'r_1(x, y) - I'r_2(x + 1, y)|$ обусловлена различием гистограмм в этих окрестностях. При значительном перекрытии окон разность мала.

Для минимизации пограничного эффекта используется взвешенное объединение:

$$I'(x, y) = \alpha(x, y) \cdot I'r_1(x, y) + (1 - \alpha(x, y)) \cdot I'r_2(x, y), \tag{1}$$

где $\alpha(x, y)$ ($0 \leq \alpha \leq 1$) – весовой коэффициент, плавно изменяющийся вблизи границы.

Динамическое окно обеспечивает локально-адаптивное контрастирование. Плавное изменение $r(x, y)$ и взвешенное объединение минимизируют артефакты на границах. Существует непрерывный переход от I' при $r_{\min} = 0$ (нет обработки) к глобальному контрастированию при r_{\max} .

Адаптивный размер окна

Выбран размер окна $r(x, y)$ в каждой точке изображения (рисунок 2).

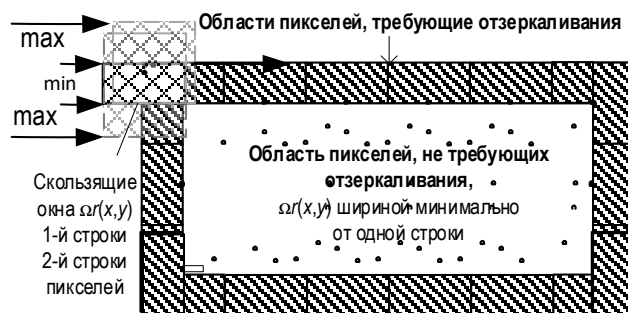


Рис. 2. Обработка изображения «скользящим окном» $\Omega r(x, y)$

Fig. 2. Processing Using the $\Omega r(x, y)$ Sliding Window

При промежуточных размерах окна возникает компромисс: меньший радиус лучше сохраняет мелкие текстуры (усиливает их контраст), но повышает шум. Предлагается использовать локальную энтропию, частотные характеристики и результаты безэталонной оценки качества как количественные показатели, определяющие размер окна. Обработка изображений производится в соответствии с алгоритмом.

Локальная энтропия

Оценка неоднородности (энтропия показывает, насколько разнообразны уровни яркости в локальной области):

$$H_{\Omega r}(x, y) = \sum_{i=0}^N p_i \log(p_i), \tag{2}$$

где p_i – относительная частота яркости i в окрестности $\Omega r(x, y)$; N – количество пикселей в локальной области; i – уровни яркости.

Из теории информации известно: мера Хартли (энтропия источника = максимальная энтропия) – частный случай меры Шеннона (энтропия сообщения) при равновероятных символах в сообщении. Локальная энтропия используется для управления параметрами не только окна, но и контрастера. Энтропийный критерий используется для оценки неоднородности яркости в локальной области изображения, что позволяет алгоритму адаптивно реагировать на изменения текстурных характеристик сцены и динамически выбирать параметры контрастирования.

Частотный анализ

На этапе частотного анализа к фрагменту изображения $I(x+i, y+j)$ размером $(2r+1) \times (2r+1)$ применяется двумерное дискретное косинусное преобразование (ДКП) для оценки его спектральных характеристик.

Первый коэффициент ДКП представляет собой среднюю яркость в окне (ДК-компонента), а коэффициенты F_{uv} (коэффициенты в разложении) при больших значениях u и v соответствуют высокочастотным компонентам изображения (мелким деталям, резким перепадам яркости, шуму).

Введем $RHF(x, y)$ как функцию оценки частот в локальной области. Доля высокочастотной информации $RHF(x, y)$ оценивается как отношение энергии высоких частот к полной энергии в окне:

$$RHF(x, y) = \frac{(\sum_{u,v=1}^r (u, v) \in H |F(u, v)|^2)}{(\sum_{u,v=1}^r (u, v) \in \Omega |F(u, v)|^2)}, \quad (3)$$

где u и v – индексы частот по горизонтали и вертикали, соответственно, и $-r \leq u, v \leq r$; Ω – множество всех пар индексов (u, v) в пределах окна ДКП, т. е. H – множество пар индексов (u, v) , соответствующих высоким частотам и определяемое пороговым значением D , которое определяет границу между низкими и высокими частотами:

$$H = \{(u, v) \mid u^2 + v^2 > D^2\};$$

$|F(u, v)|^2$ – квадрат модуля коэффициента ДКП, представляющий собой энергию соответствующей частотной компоненты.

Выбор D зависит от размера максимального окна r_{\max} обработки и максимальной частоты фона. Выбрано $D = r/2$, что соответствует половине максимальной частоты в обрабатываемом окне.

Безэталонная оценка качества изображения E вычисляется как взвешенное среднее величин искажений по всем пикселям:

$$E = \frac{(\sum_{k=1}^H K_{\text{иск}}) + C_1}{(N + C_1)}, \quad (4)$$

где E – результат оценки качества изображений; $K_{\text{иск}}$ – коэффициент искажения (ни в одном из вы-

шеперечисленных случаев не превышает значения 1); N – число активных локальных областей в изображении, на которые было произведено разбиение; C_1 – положительная константа (в рассматриваемом случае $C_1 = 1$), которая введена для предотвращения неопределенности при знаменателе переходящем в 0.

Итоговый результат оценки качества изображений определяется значениями: $0,1T < K_{\text{иск}} < T$ [4].

Итоговый результат регулировки окна:

$$r_{\text{target}}(x, y) = (1 - \beta) \times \frac{[r_{\min} + (r_{\max} - r_{\min}) \times (\eta(x, y) + \nu(x, y))]}{2} + \beta * r(x, y) \times \frac{E(x, y)}{100}, \quad (5)$$

где $\eta = \frac{(H_{\max} - H)}{H_{\max}}$ – мера энтропии; $\nu = 1 - RHF$ – частотный коэффициент; β – весовой коэффициент для общей оценки качества и количества артефактов в предыдущем кадре.

Для устранения проблем с обработкой ИК изображений, а также для сюжета с широким динамическим диапазоном, в алгоритм добавлено уменьшение размера окна. У границ изображения и в областях с разной яркостью или сюжетом $\Omega r(x, y)$ – окно уменьшается:

$$\Omega r(x, y) \leftarrow \Omega r(x, y) \times \frac{(1 - \alpha * ||\nabla I(x, y)||)}{\max_{\text{gradient}} - \text{gradient}_{\text{threshold}}, \text{ при } ||\nabla I(x, y)||} > \quad (6)$$

Ограничение контраста (задается порогом T_c) снижает коэффициент усиления до величины, пропорциональной $\frac{(L-1)T_c}{|\Omega r|}$, где $|\Omega r|$ – количество пикселей в окне. Чем меньше размер окна (и, следовательно, меньше ΔI), тем больше потенциальный коэффициент усиления шума. Коэффициент T_c выбирается в зависимости от оценки уровня шума в текущей области обработки, регулируя баланс между усилением контраста и подавлением шума:

$$|T(i) - T(j)| \leq K|i - j|, \quad (7)$$

где $K < \infty$ и определяется значением $\frac{T_c}{|\Omega r|}$.

Алгоритм обеспечивает адаптивный выбор размера окна, стремясь удержать локальную энтропию и долю высоких частот на заданном уровне, а также учитывая качество обработки на предыдущем кадре и предотвращая появление артефактов у границ.

Внедрение ИИ и использование нейросетевой модели

Ключевой особенностью разработанного метода является применение нейронных сетей для автоматической настройки параметров адаптивного

локального контрастирования (АЛК). Согласно теореме Цыбенко (универсальной аппроксимации) [11], любая непрерывная функция на компактном подмножестве конечномерного пространства может быть сколь угодно точно приближена нейронной сетью, с одним скрытым слоем, с достаточно большим числом нейронов к параметрам адаптивного локального контрастирования. Это теоретически обосновывает возможность использования нейросетей для аппроксимации нелинейных преобразований яркости и регрессии оптимальных параметров алгоритма контрастирования, что особенно важно при обработке изображений с низкой контрастностью.

Архитектура нейросетевой модели

Для реализации предложенного метода используется гибридный подход с применением CNN, состоящий из двух основных компонентов:

1) CNN для сегментации фона, основанная на архитектуре U-Net, которая благодаря своей U-образной форме с пропускными соединениями между энкодером и декодером, эффективно сочетает локальную информацию (для точной локализации границ) и глобальный контекст (для распознавания областей фона);

2) CNN для регрессии параметров АЛК.

Общая структура гибридной модели

Модель принимает на входе исходное изображение и выдает два результата:

- маску фона (вероятность принадлежности пикселей к фоновой области);
- вектор параметров АЛК (размер окна обработки и порог ограничения контраста).

Архитектура состоит из модулей (рисунок 3).

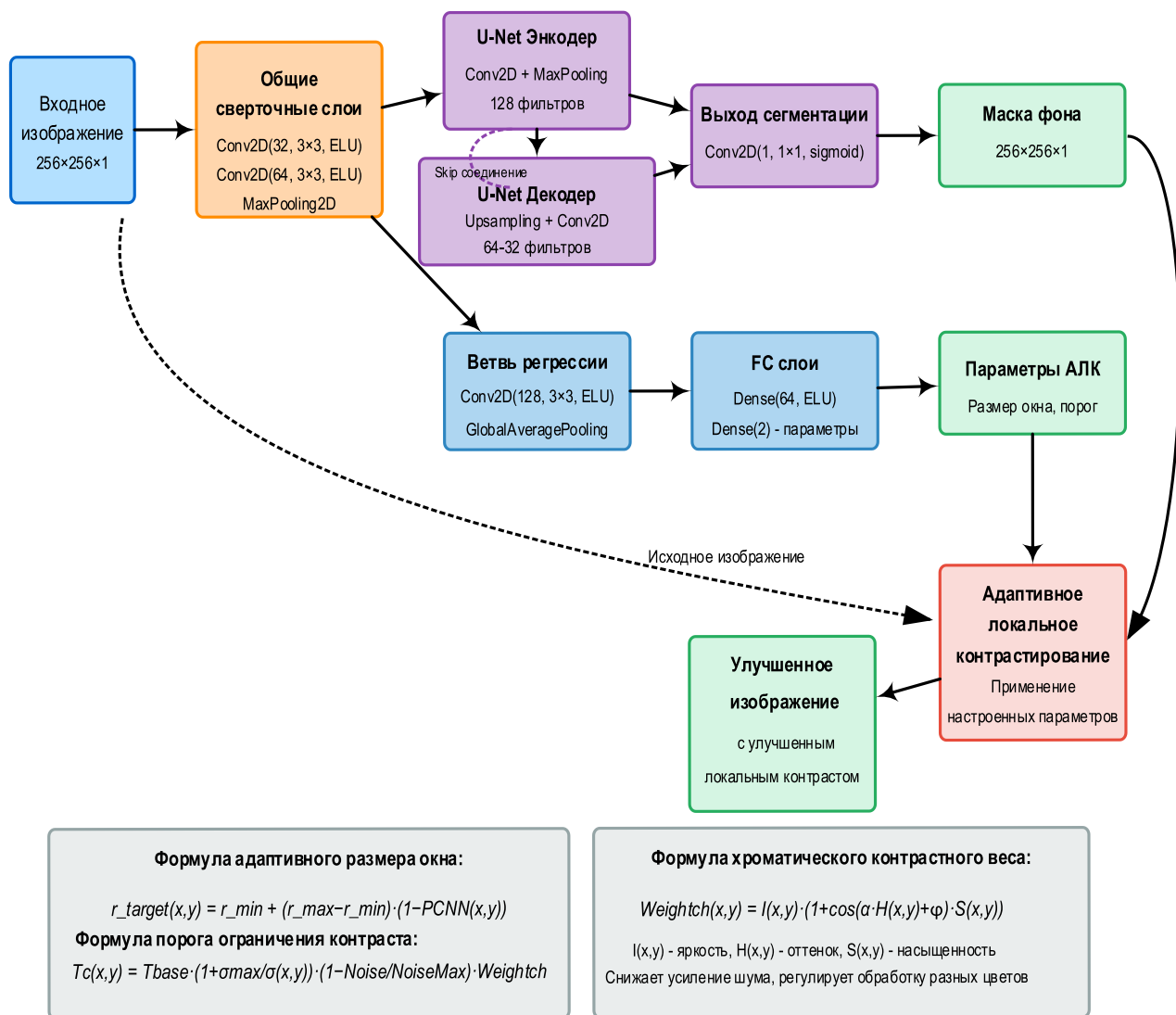


Рис. 3. Архитектура гибридной CNN для регулировки адаптивного локального контрастирования

Fig. 3. Architecture of the Hybrid Convolutional Neural Network (CNN) for Adaptive Local Contrast Enhancement Adjustment

1) Входной слой: принимает изображение в градациях серого или компоненты цветного изображения.

2) Общие сверточные слои: 3 сверточных слоев с фильтрами (количество которых составляет 32) и ядром 3×3, с функцией активации ELU для извлечения общих признаков из изображения.

3) Ветвь сегментации фона (U-Net) [13]:

– энкодер: сверточные слои и слои субдискретизации;

– декодер: сверточные слои и слои апсэмплинга;

– пропускные соединения между энкодером и декодером;

– выходной слой с сигмоидной активацией для получения маски фона.

4) Ветвь регрессии параметров:

– сверточные слои с глобальным усреднением;

– полносвязные слои;

– линейный выходной слой для предсказания параметров АЛК.

5) Skip connections соединяют соответствующие слои энкодера и декодера. На выходном слое используется сверточный слой для предсказания 4 параметров.

Обучение для видимого спектра: 150 эпох.

1) Размер батча: 99 изображений (равномерная выборка по 33 изображения из каждой подкатегории видимого спектра).

2) Оптимизатор: Стохастический градиентный спуск (SGD).

3) Начальная скорость обучения: 0,001 (скорость изменения коэффициентов слоев градиентного спуска).

4) Стратегия Learning Rate: Cosine annealing. Скорость обучения циклически изменяется по косинусоидальной функции, способствуя стабильности сходимости и предотвращению локальных минимумов. Параметры cosine annealing (период цикла, минимальная и максимальная скорость обучения) были настроены эмпирически для достижения оптимальной сходимости.

5) Функция потерь: среднее значение безэталонной оценки качества по батчу. Цель – минимизация функции потерь, т. е. улучшение качества изображения по собственной метрике безэталонной оценки качества.

CNN используется для оценки локальной контрастности изображения, определения параметров АЛК и сегментации фона. В качестве функции активации в сверточных слоях используется экспоненциальная линейная функция (ELU), которая, в отличие от ReLU, позволяет получать и отрицательные выходные значения, что полезно для представления информации о контрасте. На выходе сети формируется вектор значений полученных нейронной сетью.

Адаптивное ограничение контраста

Для предотвращения чрезмерного усиления шума применяется адаптивное ограничение контраста с порогом T_c , вычисляемым по формуле:

$$T_c(x, y) = T_{base} \cdot \left(1 + \frac{\sigma_{max}}{\sigma(x, y)}\right) \times \left(1 - \frac{Noise_{max}}{NoiseLevel(x, y)} \cdot Weightch(x, y)\right), \quad (8)$$

где T_{base} – базовый порог; $\sigma(x, y)$ – локальная дисперсия; σ_{max} – максимальная дисперсия; $Noise_{max}$ – максимальный уровень шума; $NoiseLevel(x, y)$ – оценка уровня шума; $Weightch(x, y)$ – хроматический контрастный вес.

Обучение нейросетевой модели

1) Прямой проход: Изображение -> Нейросеть -> 4 параметра -> АЛК -> Обработанное изображение.

2) Оценка качества: Обработанное изображение -> безэталонная оценка качества изображения -> Значение оценки.

3) Расчет функции потерь: Функция потерь = Значение безэталонной оценки качества изображения (среднее по батчу).

4) Обратное распространение: Вычисление градиентов функции потерь по весам CNN через backpropagation. Градиенты определяют направление изменения весов для минимизации функции потерь.

5) Обновление весов: Оптимизатор градиентный спуск обновляет веса CNN на основе градиентов для улучшения предсказания параметров АЛК, оптимизируя значение безэталонной оценки.

Обучающая выборка

Для обучения гибридной сети использовался датасет из более чем 200 000 изображений, охватывающий видимый и ИК спектры, а также различные условия съемки. Данные разделены по спектральному диапазону.

Данные видимого Спектра (3 500 пар):

– черно-белые (1920×1200): 1000 пар (солнечно / пасмурно); центральная часть изображений без сжатия;

– черно-белые (1920×1200): 1500 пар (с / без поляризационного фильтра); центральная часть изображений без сжатия;

– цветные (1280×1024 HSV используется V-канал): 1000 пар для каждого фильтра KC10 ($\lambda > 600$ нм) и KC13 ($\lambda > 650$ нм)(красные фильтры).

Данные ИК спектра (11 500 изображений): 11 500 изображений с NUC (коррекция неоднородностей матриц с двумя кривыми второго порядка); для изображений 1280×024 использовалась центральная область 640×512 (16-битные изображения).

Для ИК и видимого диапазона были получены две отличающиеся модели нейронной сети для

настройки АЛК. Обучение модели проводилось на двух видеокартах NVIDIA Tesla K40X и заняло около 200 000 итераций.

Детализированные параметры обучения разработанной нейронной сети представлены в таблице 2.

ТАБЛИЦА 2. Параметры обучения
TABLE 2. CNN Training Parameters

Параметр	Значение	Обоснование
Эпохи	100	Обеспечивает сходимость без переобучения
Размер батча	99	Сбалансированное использование GPU
Learning rate	1e-4 → 1e-6	Используется scheduler ReduceLRonPlateau
Оптимизатор	Градиентный спуск	Снижение переобучения, стабильность
Функция потерь	Комбинированная безэталоная	Универсальность, не требует разметки
Инициализация	Xavier Uniform	Оптимальна для ELU
Активация	ELU ($\alpha = 0,1$)	Предотвращает затухание градиента
Регуляризация	Dropout 0,3 (в полносвязных слоях)	Устойчивость к переобучению
Аугментации	Различие выборок входных данных	Повышение обобщающей способности
Набор данных, без меток	> 3 500 изображений видимого спектра > 11 500 ИК изображений	Аэрофотосъемка 1920×1200 / 1280×1024, ИК кадры
Фреймворк	PyTorch + Albumentations	Быстрая отладка, кастомные трансформации

Экспериментальная часть

Эффективность предложенного подхода подтверждается экспериментально с использованием объективных метрик. В качестве объективных метрик применялись:

- структурный индекс подобия (SSIM, аббр. от англ. Structural Similarity Index Measure – метрика для оценки схожести двух изображений, учитывающая структурную информацию, яркость и контраст; часто считается более близкой к человеческому восприятию, чем пиковое отношение сигнала к шуму или среднеквадратичная ошибка (MSE, аббр. от англ. Mean Squared Error): измеряет структурное сходство между двумя изображениями, принимая значения от 0 до 1, где 1 означает идеальное соответствие;

- средняя точность (mAP, аббр. от англ. mean Average Precision – средняя точность (Average Precision), усредненная по всем классам объектов или по разным порогам; стандартная метрика для оценки качества моделей обнаружения объектов (object detection): оценивает точность обнаружения объектов на изображении.

Для mAP использовались размеченные изображения:

- CityScapes (видимый спектр): 20 000 изображений 1920×1080 с людьми; оценка детекции YoLOv5 по mAP;

- FLIR (ИК спектр): 10 000 изображений 640×512 с людьми; оценка детекции YoLOv5 по mAP.

При mAP = 0,4 ошибка ϵ составляет не более 0,05 – это соответствует выборке из 20 000 изображений (в 20 раз больше 1/0,05). Ошибка ϵ , тестовая выборка $\geq 1/\epsilon$.

Для экспертной оценки качества использовался метод парного сравнения, в котором экспертам предлагалось выбрать изображение с лучшим контрастом из пары, включающей результат работы предложенного алгоритма и результат одного из следующих методов:

- CLANE: с параметрами, указанными в оригинальной работе [19];

- Multi-Scale Retinex: реализация, описанная в работе [1].

Для оценки статистической значимости различий между предложенным методом и сравниваемыми методами использовался *t*-критерий Стьюдента с уровнем значимости $p < 0,05$ [5].

Результат обработки

На рисунках 4а и 4b представлено изображение туманной улицы из набора CityScapes, а на рисунках 4с и 4d – туманной улицы автострады вечером в тумане. Обработка увеличивает объективные показатели качества для малоконтрастных изображений в условиях ограниченной видимости.

При повышении точности обнаружения mAP получено 8-кратное увеличение количества обнаружений людей, а на ИК изображении – до 40 %; точность обнаружения в оригинальном ИК изображении 5 % (таблица 3).

ТАБЛИЦА 3. Количественные оценки методов обработки
TABLE 3. Quantitative Metrics for Image Processing Methods

Метод (средняя оценка 20 000 CityScapes изображений)	SSIM	mAP	<i>E</i> (без эталона)	Экспертная оценка (по рекомендации MSE)
Оригинальные	–	0,05	0,1	0,05
CLANE	0,4	0,09	0,8	0,2
MSR	0,35	0,078	0,3	0,38
Разработанный гибридный метод	0,8	0,4	0,1	0,78

Результаты обнаружения людей на ИК изображении с обработкой изображений методом CLANE показали излишнее усиление шума из-за чего получаются ложные обнаружения (рисунок 5).

На рисунке 6 практической реализации разработанного алгоритма АЛК на ПЛИС, включая реализацию нейросетевой настройки.



Рис. 4. Результат обработки изображения: оригинальное (а, с) и обработанное гибридным алгоритмом (b, d)
 Fig. 4. Processing Result: Original Image (a, c) and Image Processed with the Hybrid Algorithm (b, d)

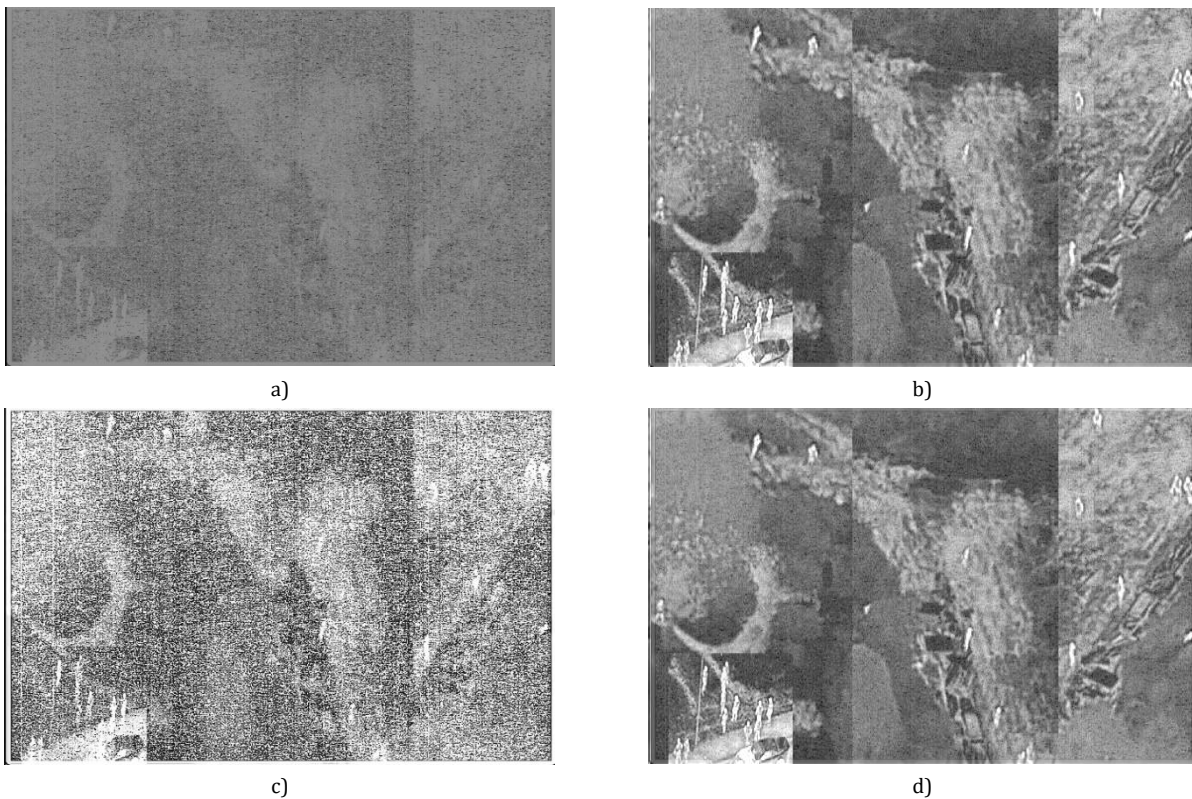


Рис. 5. Результат обработки изображения: оригинальное ИК (а), обработанное методом Clahe (с) и гибридным алгоритмом (b, d)

Fig. 5. Processing Result: Original Infrared Image (a), Image Processed with CLAHE Method (c) and Images Processed with the Hybrid Algorithm (b, d)

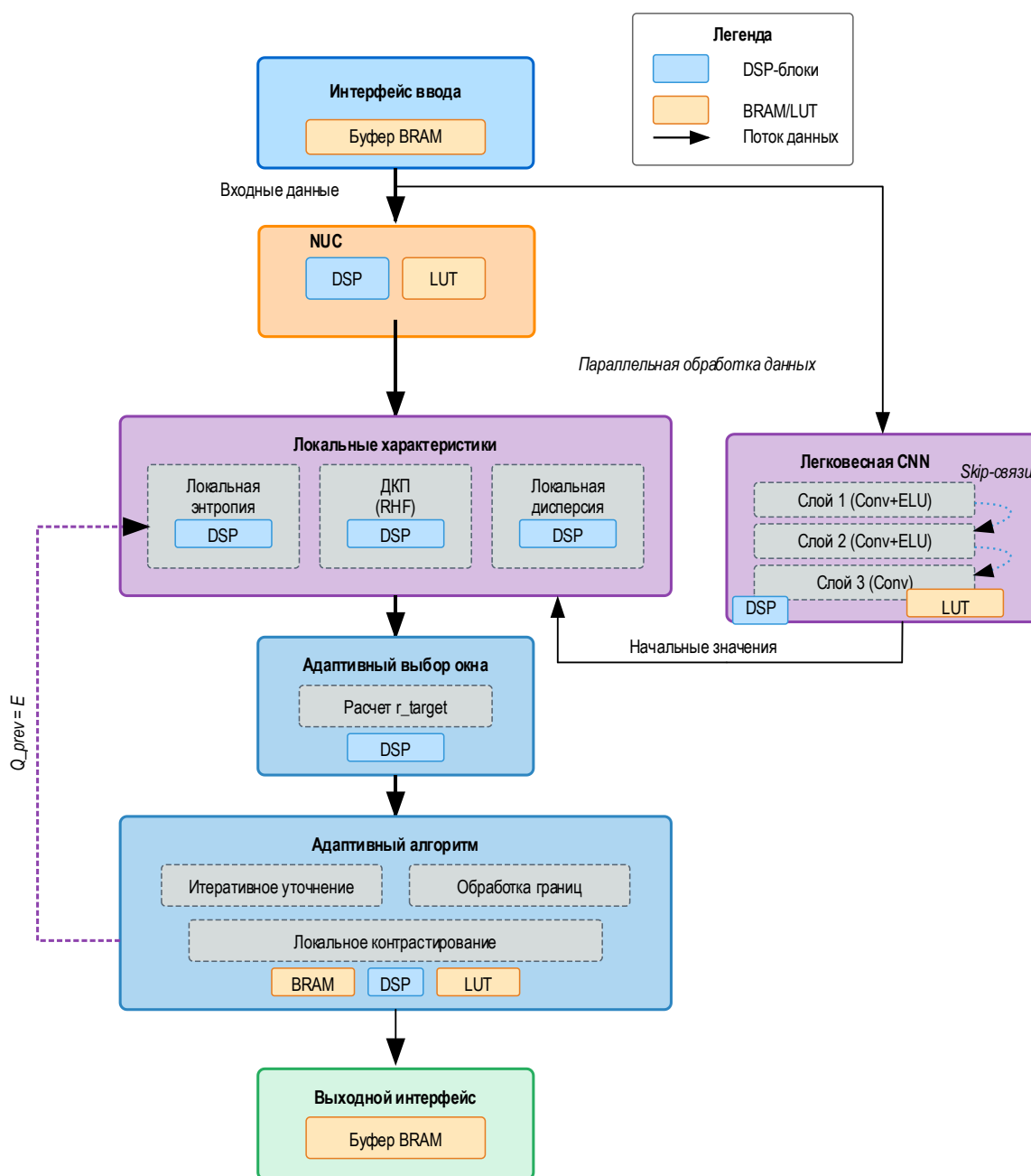


Рис. 6. Структурная схема реализации алгоритма на ПЛИС

Fig. 6. Block Diagram of the Algorithm Implementation on Field-Programmable Gate Array

Вывод

Разработанный алгоритм адаптивного локального контрастирования, основанный на динамическом определении размера окна и параметров ограничения контрастности, показал свою эффективность. Комплексный подход (АЛК + CNN) успешно протестирован на различных наборах данных, включая выборку ИК изображений (более 40 ТБ) полигонных испытаний и стандартный набор CityScapes.

Создана эффективная архитектура на ПЛИС, для реализации алгоритма с легкой нейросетью, обеспечивающая обработку в реальном времени (65 кадров/с при разрешении 1920×1080) при потреблении менее 70 % ресурсов ПЛИС и энергопотреблении менее 5 Вт.

Применение разработанного алгоритма в качестве обработки для оценки обнаружений позволило существенно увеличить (на 40 %) обнаружение людей. Обработка в системах с использованием элементов искусственного интеллекта была успешно проверена (см. рисунки 4 и 5).

Список источников

1. Jobson D.J., Rahman Z., Woodell G.A. Properties and performance of a center/surround retinex // *IEEE Transactions on Image Processing*. 1997. Vol. 6. Iss. 3. PP. 451–462. DOI:10.1109/83.557356
2. Chen Y.S., Wang Y.C., Kao M.H., Chuang Y.Y. Deep Photo Enhancer: Unpaired Learning for Image Enhancement from Photographs with GANs // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR, Salt Lake City, USA, 18–23 June 2018)*. IEEE, 2018. DOI:10.1109/CVPR.2018.00660
3. Paris S., Hasinoff S.W., Kautz J. Local Laplacian filtering: edge-aware image processing with a Laplacian pyramid // *Proceedings of the Conference on Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH '11, Vancouver, Canada, 7–11 August 2011)*. New York: Association for Computing Machinery, 2011. URL: https://people.csail.mit.edu/sparis/publi/2011/siggraph/Paris_11_Local_Laplacian_Filters_lowres.pdf (Accessed 25.04.2025)
4. Грицкевич И.Ю., Гоголь А.А. Алгоритм безэталонной оценки качества изображений // *Труды учебных заведений связи*. 2024. Т. 10. № 2. С. 16–23. DOI:10.31854/1813-324X-2024-10-2-16-23. EDN:ТТРABW
5. Rec. ITU-R BT.500-11. Methodology for subjective assessment of the quality of television pictures. ITU-T. 2002.
6. Шелепин Ю.Е. Введение в нейрокогнику. СПб.: Троицкий мост, 2017. 352 с. EDN:YNTJRJ
7. Kim Y.T. Contrast enhancement using brightness preserving bi-histogram equalization // *IEEE Transactions on Consumer Electronics*. 1997. Vol. 43. Iss. 1. DOI:10.1109/30.580378
8. Rahman Z., Jobson D.J., Woodell G.A. Multi-scale retinex for color image enhancement // *Proceedings of the 3rd International Conference on Image Processing (Lausanne, Switzerland, 19 September 1996)*. IEEE, 1996. DOI:10.1109/ICIP.1996.560995
9. Ying Z., Li G., Gao W. A Bio-Inspired Multi-Exposure Fusion Framework for Low-light Image Enhancement // *arXiv preprint arXiv:1711.00591*. 2017. DOI:10.48550/arXiv.1711.00591
10. Vala H.J., Baxi A. A review on Otsu image segmentation algorithm // *International Journal of Advanced Research in Computer Engineering & Technology*. 2013. Vol. 2. Iss. 2. PP. 387–389.
11. Cybenko G. Approximation by superpositions of a sigmoidal function // *Mathematics of Control, Signals and Systems*. 1989. Vol. 2. PP. 303–314. DOI:10.1007/bf02551274. EDN:OKSIPR
12. Wang R., Zhang Q., Fu C.W., Shen X., Zheng W.S., Jia J. Underexposed Photo Enhancement Using Deep Illumination Estimation // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR, Long Beach, USA, 15–20 June 2019)*. IEEE, 2019. DOI:10.1109/CVPR.2019.00701 (13)
13. Han Y., Ye J.C. Framing U-Net via Deep Convolutional Framelets: Application to Sparse-View CT // *IEEE Transactions on Medical Imaging*. 2018. Vol. 37. Iss. 6. PP. 1418–1429. DOI:10.1109/TMI.2018.2823768
14. What is Histogram Equalization and how it works? // *Great Learning*. 2024. URL: <https://www.mygreatlearning.com/blog/histogram-equalization-explained> (Accessed 25.04.2025)
15. Liu J., Li D., Yuan C., Luo B., Wu G. A low-light image enhancement method with brightness balance and detail preservation // *PLoS One*. 2022. Vol. 17. Iss. 5. P. e0262478. DOI:10.1371/journal.pone.0262478. EDN:DFDSOY
16. Pizer S.M., Amburn E.P., Austin J.D., Cromartie R., Geselowitz A., Greer T., et al. Adaptive histogram equalization and its variations // *Computer Vision, Graphics, and Image Processing*. 1987. Vol. 39. Iss. 3. PP. 355–368. DOI:10.1016/S0734-189X(87)80186-X
17. Kaur M., Kaur J., Kaur J. Survey of Contrast Enhancement Techniques based on Histogram Equalization // *International Journal of Advanced Computer Science and Applications*. 2011. Vol. 2. Iss. 7. DOI:10.14569/IJACSA.2011.020721
18. Zuiderveld K. VIII.5. – Contrast Limited Adaptive Histogram Equalization // In: Heckbert P.S. (ed.) *Graphics Gems IV*. Academic Press, 1994. PP. 474–485. DOI:10.1016/B978-0-12-336156-1.50061-6
19. Lore K.G., Akintayo A., Sarkar S. LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement // *Pattern Recognition*. 2017. Vol. 61. PP. 650–662. DOI:10.1016/j.patcog.2016.06.008
20. Wei C., Wang W., Yang W., Liu J. Deep Retinex Decomposition for Low-Light Enhancement. 2018. URL: http://39.96.165.147/Pub%20Files/2018/chen_bmvc18.pdf (Accessed 25.04.2025)
21. Liu X., Ma Y., Shi Z., Chen J. GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing // *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV, Seoul, South Korea, 27 October – 02 November 2019)*. IEEE, 2019. DOI:10.1109/ICCV.2019.00741
22. Guo C., Li C., Guo J., Loy C.C., Hou J., Kwong S., Cong R. Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR, Seattle, USA, 13–19 June 2020)*. IEEE, 2020. DOI:10.1109/CVPR42600.2020.00185
23. Hossain F., Alsharif M.R. Image Enhancement Based on Logarithmic Transform Coefficient and Adaptive Histogram Equalization // *Proceedings of the International Conference on Convergence Information Technology (ICCIT 2007, Gwangju, South Korea, 21–23 November 2007)*. IEEE, 2007. DOI:10.1109/ICCIT.2007.4420457
24. Stark J.A. Adaptive image contrast enhancement using generalizations of histogram equalization // *IEEE Transactions on Image Processing*. 2000. Vol. 9. Iss. 5. PP. 889–896. DOI:10.1109/83.841534
25. Потапова А.А. Новейшие методы обработки изображений. М.: ФИЗМАТЛИТ, 2008. 496 с.

References

1. Jobson D.J., Rahman Z., Woodell G.A. Properties and performance of a center/surround retinex. *IEEE Transactions on Image Processing*. 1997;6(3):451–462. DOI:10.1109/83.557356
2. Chen Y.S., Wang Y.C., Kao M.H., Chuang Y.Y. Deep Photo Enhancer: Unpaired Learning for Image Enhancement from Photographs with GANs. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 18–23 June 2018, Salt Lake City, USA*. IEEE; 2018. DOI:10.1109/CVPR.2018.00660


3. Paris S., Hasinoff S.W., Kautz J. Local Laplacian filtering: edge-aware image processing with a Laplacian pyramid. *Proceedings of the Conference on Special Interest Group on Computer Graphics and Interactive Techniques, SIGGRAPH '11, 7–11 August 2011, Vancouver, Canada*. New York: Association for Computing Machinery; 2011. URL: https://people.csail.mit.edu/sparis/publi/2011/siggraph/Paris_11_Local_Laplacian_Filters_lowres.pdf [Accessed 25.04.2025]
4. Gritskevich I., Gogol A. No-Reference Image Quality Assessment Algorithm. *Proceedings of Telecommunication Universities*. 2024;10(2):16–23. (in Russ.) DOI:10.31854/1813-324X-2024-10-2-16-23. EDN:TTPABW
5. Rec. ITU-R BT.500-11. Methodology for subjective assessment of the quality of television pictures. ITU-T. 2002.
6. Shelepin Yu.E. *Introduction to Neuroiconics*. St. Petersburg: Troickij most Publ.; 2017. 352 p. (in Russ.) EDN:YNTJRI
7. Kim Y.T. Contrast enhancement using brightness preserving bi-histogram equalization. *IEEE Transactions on Consumer Electronics*. 1997;43:1. DOI:10.1109/30.580378
8. Rahman Z., Jobson D.J., Woodell G.A. Multi-scale retinex for color image enhancement. *Proceedings of 3rd IEEE International Conference on Image Processing, 19 September 1996, Lausanne, Switzerland*. IEEE; 1996. DOI:10.1109/ICIP.1996.560995
9. Ying Z., Li G., Gao W. A Bio-Inspired Multi-Exposure Fusion Framework for Low-light Image Enhancement. *arXiv preprint arXiv:1711.00591*. 2017. DOI:10.48550/arXiv.1711.00591
10. Vala H.J., Baxi A. A review on Otsu image segmentation algorithm. *International Journal of Advanced Research in Computer Engineering & Technology*. 2013;2(2):387–389.
11. Cybenko G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*. 1989;2:303–314. DOI:10.1007/bf02551274. EDN:OKSIPR
12. Wang R., Zhang Q., Fu C.W., Shen X., Zheng W.S., Jia J. Underexposed Photo Enhancement Using Deep Illumination Estimation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 15–20 June 2019, Long Beach, USA*. IEEE; 2019. DOI:10.1109/CVPR.2019.00701
13. Han Y., Ye J.C. Framing U-Net via Deep Convolutional Framelets: Application to Sparse-View CT. *IEEE Transactions on Medical Imaging*. 2018;37(6):1418–1429. DOI:10.1109/TMI.2018.2823768
14. *Great Learning*. What is Histogram Equalization and how it works? 2024. URL: <https://www.mygreatlearning.com/blog/histogram-equalization-explained> [Accessed 25.04.2025]
15. Liu J., Li D., Yuan C., Luo B., Wu G. A low-light image enhancement method with brightness balance and detail preservation. *PLoS One*. 2022;17(5):e0262478. DOI:10.1371/journal.pone.0262478. EDN:DFDSOY
16. Pizer S.M., Amburn E.P., Austin J.D., Cromartie R., Geselowitz A., Greer T., et al. Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*. 1987;39(3):355–368 DOI:10.1016/S0734-189X(87)80186-X
17. Kaur M., Kaur J., Kaur J. Survey of Contrast Enhancement Techniques based on Histogram Equalization. *International Journal of Advanced Computer Science and Applications*. 2011;2(7). DOI:10.14569/IJACSA.2011.020721
18. Zuiderveld K. VIII.5. – Contrast Limited Adaptive Histogram Equalization. In: Heckbert P.S. (ed.) *Graphics Gems IV*. Academic Press; 1994. p.474–485. DOI:10.1016/B978-0-12-336156-1.50061-6
19. Lore K.G., Akintayo A., Sarkar S. LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement. *Pattern Recognition*. 2017;61:650–662. DOI:10.1016/j.patcog.2016.06.008
20. Wei C., Wang W., Yang W., Liu J. *Deep Retinex Decomposition for Low-Light Enhancement*. 2018. URL: http://39.96.165.147/Pub%20Files/2018/chen_bmvc18.pdf [Accessed 25.04.2025]
21. Liu X., Ma Y., Shi Z., Chen J. GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. *Proceedings of the IEEE/CVF International Conference on Computer Vision, ICCV, 27 October – 02 November 2019, Seoul, South Korea*. IEEE; 2019. DOI:10.1109/ICCV.2019.00741
22. Guo C., Li C., Guo J., Loy C.C., Hou J., Kwong S., Cong R. Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 13–19 June 2020, Seattle, USA*. IEEE; 2020. DOI:10.1109/CVPR42600.2020.00185
23. Hossain F., Alsharif M.R. Image Enhancement Based on Logarithmic Transform Coefficient and Adaptive Histogram Equalization. *Proceedings of the 2007 International Conference on Convergence Information Technology, ICCIT 2007, 21–23 November 2007, Gwangju, South Korea*. IEEE; 2007. DOI:10.1109/ICCIT.2007.4420457
24. Stark J.A. Adaptive image contrast enhancement using generalizations of histogram equalization. *IEEE Transactions on Image Processing*. 2000;9(5):889–896. DOI:10.1109/83.841534
25. Potapova A.A. *The Latest Methods of Image Processing*. Moscow: FIZMATLIT Publ.; 2008. 496 p. (in Russ.)

Статья поступила в редакцию 18.03.2025; одобрена после рецензирования 14.04.2025; принята к публикации 21.04.2025.

The article was submitted 18.03.2025; approved after reviewing 14.04.2025; accepted for publication 21.04.2025.

Информация об авторе:

**ГРИЦКЕВИЧ
Иван Юрьевич**

аспирант кафедры телевидения и метрологии Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0009-0007-4916-1816>

Автор сообщает об отсутствии конфликтов интересов.

The author declares no conflicts of interests.