

13. Al-Azawi R.J. Enhancing recognition accuracy and efficiency through intelligent frame selection in uncontrolled conditions, *J Opt.*, 2024, pp. 1-7. DOI: <https://doi.org/10.1007/s12596-024-02191-4>.
14. Zhao L. A Hybrid deep learning-based intelligent system for sports action recognition via visual knowledge discovery, *IEEE Acces.*, 2023, Vol. 11, pp. 46541-46549. DOI: 10.1109/ACCESS.2023.3275012.
15. Zhou J., Jiangbo Ai. Zheng Wang, Chen S., Wei Q. Discovering attractive segments in the user generated video streams, *Lecture Notes in Computer Science*, 2019, Vol. 11642, pp. 236-250. DOI: https://doi.org/10.1007/978-3-030-26075-0_18.
16. Mingju Chen, Han Xiaofeng, Zhang Hua, Lin Guojun, Kamruzzaman M.M. Quality-guided key frames selection from video stream based on object detection, *Journal of Visual Communication and Image Representation*, 2019, Vol. 65, pp. 1-7. DOI: <https://doi.org/10.1016/j.jvcir.2019.102678>.
17. Gharabagh Abdorreza, Hajhashemi Vahid, Ferreira Marta, Machado José, Tavares Joao Best Frame Selection to Enhance Training Step Efficiency in Video-Based Human Action Recognition, *Applied Sciences*, 2022, Vol. 12, pp. 1-12. DOI: 10.3390/app12041830.
18. Trofimenko Ya.M., Vinogradova L.N., Ershov E.V. Algoritmy predobrabotki izobrazheniy v sisteme identifikatsii lits v videopotoke [Algorithms for image preprocessing in the face identification system in the video stream], *Vestnik Cherepovetskogo gosudarstvennogo universiteta* [Bulletin of the Cherepovets State University], 2019, No. 4, pp. 21-29. DOI: 10.23859/1994-0637-2019-4-91-2.
19. Kumari P., Reddy S.R.N., Yadav R. Application of Dempster-Shafer Theory in Sensor Data Fusion, *Proceedings of the International Conference on Machine Learning, Deep Learning and Computational Intelligence for Wireless Communication*, 2024, pp. 91-102. DOI: https://doi.org/10.1007/978-3-031-47942-7_9.
20. Recognition of human actions. Available at: <https://www.csc.kth.se/cvap/actions/> (accessed 24 February 2026).
21. Computer-aided identification of degenerative neuromuscular diseases based on gait dynamics and ensemble decision tree classifiers. Access mode: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0252380> (accessed 24 February 2026).
22. Zhang Z. A gait recognition method for a moving target image in sports based on a decision tree, *International Journal of Biometrics*, 2021, Vol. 13, No. 2, pp. 165-179. DOI: 10.1504/IJBM.2021.114642.

Колоденкова Анна Евгеньевна – Самарский национальный исследовательский университет им. акад. С.П. Королева (Самарский университет); e-mail: anna82_42@mail.ru; г. Самара, Россия; д.т.н.; доцент; профессор кафедры «Информационные системы и технологии».

Бочкарев Михаил Олегович – Самарский национальный исследовательский университет им. акад. С.П. Королева (Самарский университет); e-mail: mikhail19997@gmail.com; г. Самара, Россия; ассистент кафедры «Информационные системы и технологии».

Kolodenkova Anna Evgenievna – Samara National Research University named after Academician S.P. Korolev (Samara University); e-mail: anna82_42@mail.ru; Samara, Russia; dr. of eng. sc.; associate professor; professor of the Department of Information Systems and Technologies.

Bochkarev Mikhail Olegovich – Samara National Research University named after Academician S.P. Korolev (Samara University); e-mail: mikhail19997@gmail.com; Samara, Russia; assistant of the Department of Information Systems and Technologies.

УДК 004.4

DOI 10.18522/2311-3103-2026-3-113-133

Д.А. Морозов, В.В. Гилка, А.С. Кузнецова

СОВРЕМЕННЫЕ ПОДХОДЫ РАСПОЗНАВАНИЯ ЛИЦ В УСЛОВИЯХ НИЗКОЙ ОСВЕЩЁННОСТИ: ОБЗОР И КОНЦЕПЦИЯ ГИБРИДНОЙ END-TO-END АРХИТЕКТУРЫ

Рассматривается проблема надёжного распознавания лиц в таких критически важных областях, как видеонаблюдение и биометрическая аутентификация в условиях недостаточной освещённости. Существующие подходы, как правило, разделяют задачи повышения качества изображения и собственно идентификации, что приводит к накоплению ошибок и потере информативных признаков. Целью работы является преодоление этого ограничения путём разработки и теоретического обоснования гибридной end-to-end архитектуры, в которой задачи улучшения изображения и распознавания лица решаются совместно. В исследовании проводится системати-

ческий обзор современных методов, включая классические алгоритмы (выравнивание гистограмм, подавление шумов) и передовые глубокие нейронные сети (такие как EnlightenGAN, Zero-DCE, ArcFace, RetinaFace). В качестве основного предложения представлена интеграция генеративных и идентификационных модулей в единый вычислительный граф. Ключевым результатом исследования является демонстрация того, что совместная оптимизация всех этапов обработки в рамках единой модели, в отличие от разрозненных решений, принципиально меняет подход к проблеме. Теоретический анализ и сравнительная оценка концепций показывают, что предложенная архитектура обеспечивает более эффективный поток градиентов в процессе обучения, что ведёт к формированию более качественных и устойчивых к шуму признаков для идентификации. Показано, что такой подход позволяет избежать накопления ошибки между этапами и минимизировать потери информации. Новизна работы заключается в целостном, end-to-end взгляде на проблему распознавания в условиях низкой освещённости. Практическая ценность подтверждается применимостью архитектуры в реальных системах, где её внедрение потенциально позволит повысить надёжность и скорость работы за счёт объединения разнородных задач в единый оптимизируемый контур.

Распознавание лиц; низкая освещённость; видеонаблюдение; биометрическая аутентификация; улучшение изображений; гибридная архитектура; end-to-end архитектура; генеративные модели; идентификационные модули; оценка качества распознавания; Zero-DCE; ArcFace; RetinaFace; совместная оптимизация; устойчивость к шуму.

D.A. Morozov, V.V. Gilka, A.S. Kuznetsova

MODERN APPROACHES TO FACE RECOGNITION IN LOW-LIGHT CONDITIONS: A REVIEW AND THE CONCEPT OF A HYBRID END-TO-END ARCHITECTURE

The article addresses the problem of reliable face recognition in critical areas such as video surveillance and biometric authentication under low-light conditions. Existing approaches typically separate the tasks of image enhancement and face identification, which leads to error accumulation and loss of informative features. The aim of this work is to overcome this limitation by developing and theoretically substantiating a hybrid end-to-end architecture in which image enhancement and face recognition are solved jointly. The study provides a systematic review of modern methods, including classical algorithms (such as histogram equalization and noise suppression) and advanced deep neural networks (including EnlightenGAN, Zero-DCE, ArcFace, and RetinaFace). The main contribution is the integration of generative and identification modules into a single computational graph. The key result of the study is the demonstration that joint optimization of all processing stages within a unified model, unlike fragmented solutions, fundamentally changes the approach to the problem. Theoretical analysis and comparative evaluation of existing concepts show that the proposed architecture ensures a more efficient gradient flow during training, leading to the formation of higher-quality and noise-robust identity features. It is shown that this approach prevents error accumulation between stages and minimizes information loss. The novelty of the work lies in the holistic, end-to-end view of the face recognition problem under low-light conditions. The practical significance is confirmed by the applicability of the architecture in real systems, where its implementation can potentially improve reliability and processing speed by combining heterogeneous tasks into a single optimizable framework.

Face recognition; low-light conditions; video surveillance; biometric authentication; image enhancement; hybrid architecture; end-to-end architecture; generative models; identification modules; recognition accuracy assessment; Zero-DCE; ArcFace; RetinaFace; joint optimization; noise robustness.

Введение. Распознавание лиц остаётся одной из ключевых задач компьютерного зрения, находящей широкое применение в системах видеонаблюдения, биометрической аутентификации, контроле доступа и других сферах, где требуется надёжная идентификация личности по изображению. При этом на практике такие системы часто могут использоваться в условиях недостаточной освещённости – ночью, в помещениях с ограниченным светом, при съёмке с мобильных или уличных камер в плохую погоду [1]. Именно такие условия значительно снижают эффективность распознавания: ухудшается контраст, усиливается шум, искажаются ключевые черты лица. Даже самые современные нейросетевые алгоритмы, показывающие высокую точность на качественных изображениях, могут сталкиваться серьёзными ошибками при слабом освещении. Проблема распознавания лиц на тёмных изображениях активно исследуется последние годы, и за это время было предложено множество подходов к её решению – от классических методов обработки изображений до

глубинных нейросетевых архитектур. Тем не менее, несмотря на значительные успехи, универсального решения до сих пор не существует. Часть моделей ориентированы на предварительное улучшение изображения (осветление, подавление шумов), другие сосредоточены на устойчивом распознавании, игнорируя искажения освещённости. Многие системы используют каскадные схемы, в которых модули работают поэтапно и независимо друг от друга, что снижает общую точность и устойчивость к ошибкам [2].

В рамках данной статьи будет представлен обзор существующих подходов распознавания лиц в условиях низкой освещённости, систематизированы ключевые подходы и их ограничения, а также на основе проведённого анализа обоснована необходимость новой гибридной архитектуры, в которой модули нормализации, детекции и идентификации интегрированы в единый end-to-end вычислительный граф. Предлагаемая концепция базируется на современных достижениях в области генеративных моделей и сквозного обучения, а также учитывает критические аспекты устойчивости к шуму и сохранения идентичности лица.

Проблематика распознавания лиц в условиях низкого освещения. Освещённость является одним из ключевых факторов, определяющих качество изображений в задачах компьютерного зрения. В условиях слабого или неравномерного освещения снижается контрастность изображения, теряются текстурные детали, усиливаются шумы, а ключевые черты лица такие как глаза, рот, контур носа – становятся плохо различимыми или вовсе нераспознаваемыми. Для человека эти изменения могут быть компенсированы зрительным опытом и контекстом, однако для автоматических алгоритмов подобные искажения представляют серьёзную проблему. Типичные ошибки систем распознавания в таких условиях включают как полное отсутствие детекции лица, так и неправильную идентификацию – особенно в тех случаях, когда лицо частично затемнено или засвечено, а фон содержит источники избыточного света [1]. Традиционные алгоритмы, продвинутые нейросетевые модели, обучаются преимущественно на изображениях с хорошим освещением, и потому теряют устойчивость при смене условий съёмки. Это особенно критично для задач, где система не может контролировать условия съёмки: камеры видеонаблюдения в тёмное время суток, съёмка в помещении без вспомогательного света, биометрические терминалы в нестабильной среде. Также, слабая освещённость часто сопровождается увеличением сенсорных шумов – как яркостных, так и цветовых. Эти шумы усиливаются при последующей программной коррекции яркости, что ещё сильнее искажает изображение. Как следствие, возникает дилемма: с одной стороны, нужно повысить яркость, чтобы лицо стало различимым; с другой – это приводит к усилению искажений, затрудняющих распознавание. Дополнительную сложность представляет то, что уровень освещённости может сильно варьироваться по сцене: одна часть лица может быть в тени, другая – пересвечена [3]. Такие неоднородности мешают единообразной обработке изображения и требуют адаптивных решений. Также стоит отметить, что при слабом освещении часто снижается надёжность локализации ключевых точек лица, что критично для последующих этапов идентификации, особенно если используется нормализация по положению или анализ регионов интереса (глаз, рта и т.д.).

Суммируя изложенное, можно утверждать, что успешное распознавание лиц в условиях плохой освещённости требует комплексной методологии, включающей корректное усиление изображения, подавление шумов, сохранение биометрических признаков и устойчивую работу алгоритмов локализации. Только при выполнении этих условий возможно достижение стабильной точности и надёжности в системах, функционирующих в реальных эксплуатационных условиях.

Метрики оценки качества изображений при низкой освещённости. Для объективного анализа эффективности методов улучшения изображений в условиях низкой освещённости необходима система количественных метрик, отражающих как физические характеристики освещения, так и визуальное восприятие результата. Без таких метрик невозможно корректно сравнивать различные алгоритмы, поскольку субъективная оценка не всегда коррелирует с реальным качеством реконструкции деталей и распределением яркости. В современных исследованиях [4–11] используется комплекс показателей, позволяющих оценивать степень затемнённости, равномерность освещения, контрастность, сохранение структуры и естественность изображения.



Рис. 1. Распознавания лиц в условиях низкого освещения изображений

Основу количественного анализа составляют цветовые пространства, разделяющие яркостную и цветовую компоненты. Наиболее часто применяется пространство YCbCr, где компонент Y отражает яркость (luminance), а компоненты Cb и Cr отвечают за цветовые отклонения.

Преобразование RGB в YCbCr задаётся уравнениями:

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B, \\ Cb &= 0.564(B - Y), \\ Cr &= 0.713(R - Y). \end{aligned} \quad (1)$$

Это представление позволяет анализировать освещённость изображения независимо от цветовых характеристик сцены – именно компонент Y чаще используется для расчёта яркостных метрик.

Для более точного соответствия человеческому восприятию яркости также применяется перцептивное цветовое пространство CIELAB. В нём координата L^* (светлота) линейно связана с субъективным ощущением яркости и вычисляется по формуле

$$L^* = 116 f\left(\frac{Y}{Y_n}\right) - 16, \quad (2)$$

где Y – яркость белой точки.

Это пространство позволяет эффективно сравнивать изображения, полученные в разных условиях съёмки, используя для сопоставлений единую яркостную компоненту [5].

EXP (Error of Exposure) – характеризует отклонение средней яркости изображения от целевого уровня. Средняя яркость $E[Y]$ вычисляется как усреднённое значение яркости всех пикселей изображения, а целевая яркость Y^* обычно устанавливается в диапазоне 0.45–0.55, что соответствует средне-серому тону. Ошибка экспозиции:

$$EXP = |E[Y] - Y^*|. \quad (3)$$

Помимо этого, для комплексной оценки учитываются доли недоэкспонированных p_{low} и переэкспонированных p_{high} областей, где p_{low} представляет собой вероятность того, что яркость пикселя опустится ниже 0,1, а p_{high} – вероятность превышения яркостью значения 0.9.

Считается, что изображение имеет корректную освещённость, если ошибка если ошибка EXP не превышает 0.1, при этом менее 20% пикселей являются недоэкспонированными и менее 10% переэкспонированными [4, 5, 10].

Энтропия яркости $H(Y)$ представляет собой один из ключевых показателей, применяемых для количественной оценки информационной насыщенности изображения. Данная метрика отражает степень разнообразия распределения яркостных уровней и, следовательно, характеризует количество визуально различимых деталей сцены. Чем выше значение энтропии, тем богаче структура изображения и тем больше информации содержится в яркостных переходах. С математической точки зрения энтропия определяется выражением:

$$H(Y) = -\sum_{i=1}^n p_i \log_2 p_i, \quad (4)$$

где p_i обозначает вероятность появления пикселя с яркостью i .

При равномерном распределении яркостей значение энтропии достигает максимума, что соответствует хорошо освещённым и детализированным изображениям. Напротив, низкие значения $H(Y)$ указывают на преобладание однотипных тонов, когда изображение содержит мало различий между пикселями, а значит – теряет часть информации о структуре сцены. В контексте анализа изображений при низкой освещённости энтропия служит индикатором восстановления визуальной информации после применения алгоритмов улучшения. Увеличение значения $H(Y)$ после обработки свидетельствует о повышении информативности изображения и более равномерном распределении яркости. Однако чрезмерный рост энтропии может указывать на переусиление контраста или усиление шумов, что приводит к искусственному увеличению разнообразия тонов. Практически установлено, что для оптимально экспонированных изображений энтропия яркости находится в диапазоне 6–7 бит, тогда как для тёмных кадров она обычно не превышает 4 бит. Таким образом, показатель $H(Y)$ позволяет не только оценить степень потери информации в исходных данных, но и объективно сравнивать эффективность различных методов коррекции освещённости. Использование энтропии в совокупности с другими метриками – такими как динамический диапазон, контраст и показатель LOE – обеспечивает комплексный анализ визуального качества изображения и способствует более точной интерпретации результатов восстановления сцены [5, 7, 8].

Также оценивается динамический диапазон:

$$\Delta = P_{99}(Y) - P_1(Y) \quad (5)$$

вычисляемый как разница между 99-м и 1-м процентилем яркости, служит количественной мерой способности камеры одновременно передавать детали в самых тёмных и самых светлых участках сцены. Эта характеристика критически важна для создания реалистичного изображения, поскольку она определяет, насколько полно может быть отображён существующий в реальности контраст между тенями и светами. Широкий динамический диапазон позволяет избежать потери информации в сложных световых условиях, например, при контровом освещении, когда без него теневые области превращаются в сплошное чёрное пятно, а светлые – в белое. На практике это напрямую влияет на гибкость при постобработке, давая возможность восстановления скрытых деталей в тенях и светах из исходного материала, особенно при съёмке в RAW. Таким образом, данный показатель является не только технической характеристикой для сравнения камер, но и фундаментальным параметром, определяющим глубину и натуральность конечного изображения.

Метрика LOE (Lightness Order Error) используется для количественной оценки степени сохранения относительных яркостных соотношений между пикселями исходного и обработанного изображений. Её основная идея заключается в том, что корректное улучшение изображения должно восстанавливать освещённость сцены, не нарушая естественный порядок «светлее–темнее» между различными участками кадра. Иными словами, если в оригинальном изображении один объект был темнее другого, то после обработки это соотношение должно сохраняться [4].

$$LOE = \frac{1}{m} \sum_x RD(x), \quad (6)$$

где m – общее количество пикселей, $RD(x)$ – число элементов, для которых нарушилось относительное упорядочение яркости.

Высокие значения метрики указывают на сильное искажение светотеневого баланса, в то время как низкие значения свидетельствуют о сохранении естественного восприятия освещения.

Интерпретация результатов расчёта LOE имеет важное практическое значение.

Значения метрики ниже 1000 считаются признаком корректной реконструкции яркостного распределения, при котором алгоритм улучшает видимость деталей без искусственного изменения контраста и структуры освещённости. Напротив, превышение порога в 5000 обычно сопровождается появлением неестественных эффектов – таких как локальные пересветы, избыточная яркость или изменение относительных отношений между светлыми и тёмными областями. Подобные искажения делают изображение визуально менее реалистичным и снижают его пригодность для последующего анализа или распознавания.

В современных исследованиях метрика LOE используется как один из ключевых инструментов объективного сравнения алгоритмов коррекции освещённости. В частности, такие модели, как EnlightenGAN, Retinex-Net и Zero-DCE++, демонстрируют значения LOE в диапазоне 400-800, что указывает на способность этих архитектур эффективно улучшать изображения при сохранении естественного светового баланса [6, 7, 11]. Таким образом, LOE выступает не только как числовой показатель, но и как критерий визуальной достоверности результата, отражающий, насколько гармонично алгоритм вмешивается в структуру освещённости сцены [4].

Метрики AMBE (Absolute Mean Brightness Error) и EME (Enhancement Measure Estimation) применяются для комплексной оценки качества изображений после улучшения освещённости и контраста. Эти показатели позволяют количественно описать, насколько обработанное изображение сохраняет естественный уровень яркости сцены и насколько эффективно усиливается локальная детализация. AMBE оценивает отклонение средней яркости между исходным и улучшенным изображениями. Она определяется как:

$$AMBE = |E[X] - E[Y]|, \quad (7)$$

где X – исходное изображение, Y – улучшенное [8–10].

Малые значения AMBE свидетельствуют о корректной работе алгоритма: обработка усиливает детали и контраст без сдвига общей освещённости сцены. При этом значения $AMBE < 10$ указывают на хорошее соответствие между изображениями, а $AMBE < 5$ – на практически идеальную яркостную согласованность. Если же AMBE значительно возрастает, это говорит о смещении яркости, когда изображение становится слишком светлым или, наоборот, чрезмерно затемнённым, что нарушает естественное восприятие сцены и может ухудшать качество распознавания.

В отличие от AMBE, метрика EME (Enhancement Measure Estimation) предназначена для анализа локального контраста. Она позволяет определить, насколько хорошо алгоритм усиливает различия между соседними участками изображения [8, 9], которая рассчитывается по формуле:

$$EME = \frac{1}{K} \sum_{i=1}^K 20 \log_{10} \left(\frac{I_{\max}^{(i)}}{I_{\min}^{(i)}} \right), \quad (8)$$

где K – число блоков, а $I_{\max}$ и $I_{\min}$ – максимальные и минимальные значения яркости внутри каждого блока.

Интерпретация результатов основывается на том, что низкие значения $EME < 15$ характерны для изображений с недостаточным контрастом и слабой детализацией. Значения в диапазоне 25-35 соответствуют визуально комфортным изображениям, где баланс между чёткостью и естественностью достигается оптимально. Уровень около 40 обычно фиксируется у передовых моделей, демонстрирующих результаты, близкие к современному состоянию технологий (SOTA). Совместное использование AMBE и EME позволяет объективно оценить, насколько алгоритм улучшения освещённости одновременно сохраняет глобальную яркость сцены и усиливает локальные детали без появления артефактов. Эти метрики широко применяются в научных исследованиях для сравнения эффективности методов, основанных на принципах Retinex, адаптивного контрастного выравнивания и генеративных моделей [8–10].

Показатель дисперсии карты освещённости σ_{illum} применяется для оценки степени равномерности распределения света по изображению и позволяет количественно охарактеризовать наличие теней, бликов и перепадов яркости. Этот параметр базируется на модели Retinex-разложения, согласно которой изображение $I(x)$ представляется как произведение отражательной способности поверхности $R(x)$ и освещённости $L(x)$ [5, 7, 10] – $I(x) = R(x) \cdot L(x)$.

Компонент $L(x)$ отражает распределение освещения в сцене, а её статистическая вариативность описывается через дисперсию σ_{illum} .

Физический смысл данной метрики заключается в том, что освещённость в реальных сценах должна изменяться плавно и согласованно. Если же дисперсия карты освещённости высока, это свидетельствует о наличии резких переходов яркости – например, участков пересвета или глубоких теней, возникающих вследствие неравномерного освещения.

Низкие значения $\sigma_{illum} < 0.15$ характерны для кадров с равномерным освещением, где контраст между областями минимален, а распределение энергии света близко к однородному. При умеренных перепадах яркости показатель находится в диапазоне $0.15 \leq \sigma_{illum} \leq 0.3$, что указывает на естественные вариации освещения, допустимые при съёмке в реальных условиях.

Значения $\sigma_{illum} > 0.5$ интерпретируются как признак выраженной неравномерности освещения, сопровождающейся сильными тенями, бликами и локальными потерями деталей.

Использование дисперсии карты освещённости особенно важно при анализе и сравнении алгоритмов улучшения изображений, поскольку позволяет объективно оценить, насколько корректно модель восстанавливает равномерность освещения без искусственного усиления контрастов.

В задачах обработки изображений при низкой освещённости данная метрика служит надёжным индикатором того, достигается ли баланс между усилением видимости деталей и сохранением естественного распределения света в сцене.

Современные подходы, основанные на принципах Retinex и генеративных моделях, стремятся минимизировать σ_{illum} , обеспечивая тем самым визуальную согласованность и физически обоснованную реконструкцию освещения [5, 7, 10].

Для обеспечения объективного анализа необходимо использовать согласованную систему количественных метрик, охватывающих как физические характеристики освещения, так и субъективное восприятие визуального результата. Каждая из таких метрик отражает определённый аспект качества изображения: от точности передачи яркости и равномерности освещённости до степени сохранения структуры и контраста. В научных исследованиях принято выделять несколько ключевых показателей, образующих основу объективной оценки визуального качества. Сопоставление эталонных значений этих метрик, позволяет установить ориентиры для оценки эффективности алгоритмов обработки изображений. В табл. 1 представлены усреднённые диапазоны показателей, соответствующие различным уровням качества – от неудовлетворительного до уровня современных передовых решений (SOTA).

Таблица 1

Пороговые значения метрик для задач улучшения изображений

Метрика	Неудовлетворительное качество	Удовлетворительное качество	SOTA
EXP	> 0.25	< 0.1	< 0.05
$H(Y)$	< 4	6-7	> 7
LOE	> 5000	< 2000	< 800
AMBE	> 30	< 10	< 5
EME	< 15	25-35	≈ 40
σ_{illum}	> 0.5	0.15-0.3	< 0.15

Как видно, EXP и σ_{illum} характеризуют физические свойства освещения и равномерность светораспределения, тогда как $H(Y)$, LOE, AMBE и EME отражают структурное и перцептивное качество изображения. В совокупности эти метрики формируют универсальную систему оценки, используемую в научных и прикладных исследованиях для сравнения методов улучшения изображений при низкой освещённости.

Количественная оценка качества изображений, полученных в условиях недостаточной освещённости, требует учёта как объективных физических параметров сцены, так и особенностей их визуального восприятия. Для формирования комплексной метрики целесообразно сочетать показатели, характеризующие уровень и распределение света, экспози-

цию и стандартное отклонение освещённости, с метриками, оценивающими естественность цветопередачи, сохранение яркостного динамического диапазона и визуальную контрастность. Подобный интегральный подход создаёт объективную основу для сравнительного анализа алгоритмов улучшения изображений, что является критически важным в таких прикладных областях, как системы видеонаблюдения, автономные транспортные средства, цифровая реставрация фотоматериалов и биометрическая идентификация.

Методы оценки качества детекции и распознавания лиц. Корректная оценка качества систем обнаружения и распознавания лиц является необходимым условием их практической надёжности. Без строгих и воспроизводимых метрик невозможно объективно сравнивать методы и делать выводы об их пригодности для реальных применений – от видеонаблюдения до биометрической идентификации. При этом детекция и распознавание требуют различных критериев: в первом случае важна геометрическая точность выделения лиц, во втором – устойчивость идентификационных признаков при изменении позы, освещённости и выражения. Ключевой задачей этапа детекции является локализация области изображения, содержащей лицо. Для этого используется ограничивающая рамка “bounding box” представленный на рис. 2, позволяющая точно определить положение и размеры объекта. Bounding box служит базовым элементом для всех последующих операций анализа – распознавания личности, трекинга, построения масок и сегментации. Корректная локализация обеспечивает стандартизированное входное пространство для нейронных сетей и повышает устойчивость вычислений. Даже небольшое смещение границ рамки приводит к снижению точности идентификации, поэтому метрики локализации являются критически важными для всей системы [12].



Рис. 2. Ограничивающая рамка (Bounding box)

Для количественной оценки совпадения предсказанной рамки B с эталонной рамкой G применяется метрика пересечения-объединения (Intersection over Union, IoU), определяемая формулой:

$$\text{IoU}(B, G) = \frac{|B \cap G|}{|B \cup G|}. \quad (9)$$

Детекция считается корректной при $\text{IoU}(B, G) \geq \tau$,

где τ традиционно принимается равным – 0.5 для задач детекции лиц.

А для более жёсткой оценки локализации используют усреднение по нескольким порогам, как это принято в современных задачах общего детектирования объектов.

Качество модели агрегируется с использованием кривой Precision–Recall, интеграл под которой определяет показатель Average Precision (AP) усреднение AP по классам или по поднаборам сложности даёт mAP.

В WIDER FACE официально используется AP при $\tau = 0.5$, и стратифицируются на категории «Easy/Medium/Hard» по факторам масштаба поз и степени окклюзий. Таким образом, бенчмарк «заземляет» метрику на реальные трудности сцены. Для задач шире лиц применяют также метрику COCO, среднюю AP сетке порогов IoU в диапазоне 0.50 : 0.95 с шагом 0.05, что позволяет более строго учитывать точность локализации и штрафовать

избыточные детекции. В обоих случаях необходимы детальные правила сопоставления предсказаний с разметкой, устранения дублей (*Non-Maximum Suppression*, NMS), и скоринга по всем порогам – именно эти правила обеспечивают воспроизводимость и сопоставимость результатов между статьями [13]. Далее, по количеству корректных (*True Positive*, TP), ложных (*False Positive*, FP) и пропущенных (*False Positive*, FN) срабатываний вычисляются показатели точности (*Precision*) и полноты (*Recall*):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (10)$$

Эти метрики отражают компромисс между избыточными и пропущенными детекциями, что особенно важно для систем, работающих в условиях ограниченного времени отклика или при высокой плотности лиц в кадре [13].

Для интегральной оценки используется средняя точность (*Average Precision*, AP), вычисляемая как площадь под кривой *Precision-Recall*:

$$\text{AP} = \int_0^1 \text{Precision}(\text{Recall})d(\text{Recall}). \quad (11)$$

Бенчмарк WIDER FACE [12], содержащий 32 203 изображения и 393 703 размеченных лица, установил стандарт оценки AP при IoU = 0.5, разделив результаты на уровни сложности – Easy, Medium и Hard. Такая стратификация выявила слабые стороны современных детекторов: резкое падение точности при малых размерах лиц и сильной окклюзии. Аналогично, набор Fddb [13], включающий 5 171 лицо на 2 845 изображениях, использует ROC-подход, отображающий зависимость доли истинно положительных срабатываний (TRP) от доли ложных (FPR). Значения (TPR) при фиксированных (FPR) (например, 1% или 0.1%) демонстрируют устойчивость алгоритма к ложным тревогам, что особенно важно для практических сценариев с высокой ценой ошибки.

На этапе распознавания лиц применяются иные критерии, ориентированные на надёжность различения личностей. В задачах верификации (1:1) используются показатели (*True Acceptance Rate* TAR) и (*False Acceptance Rate* FAR), отражающие долю корректных и ложных приёмов при заданном пороге; точка их равенства определяет (EER) (*Equal Error Rate*). В задачах идентификации (1:N) измеряется вероятность того, что правильный ответ окажется среди первых k кандидатов (*Rank-k*), а также *TPIR@FPiR* (*True Positive Identification Rate* при фиксированном уровне *False Positive Identification Rate*). Эти метрики стандартизированы в серии бенчмарков IJB-A/B/C [14], разработанных в рамках программы IARPA Janus, и являются современным эталоном для оценки систем распознавания лиц.

В табл. 2 представлены наиболее распространённые показатели, применяемые в международных бенчмарках WIDER FACE, Fddb и IJB-C [12–14].

Комплексное использование IoU, *Precision*, *Recall*, AP, EER и *Rank-k* обеспечивает многоаспектное понимание работы модели. Высокие значения AP отражают точность локализации и устойчивость к пропускам, тогда как низкий EER указывает на надёжность идентификации. На практике важно оценивать модели не только по средним значениям, но и по устойчивости к различным факторам: размеру лица, углу наклона, степени окклюзии и условиям освещения. Исследования на WIDER FACE и IJB-C подтверждают, что такие ковариаты существенно влияют на итоговую точность, а специализированные подмножества данных («low-light», «occluded») позволяют объективно оценивать robustness систем в реальных сценариях эксплуатации [12–14]. Современная практика оценки алгоритмов детекции и распознавания лиц базируется на международных стандартах, обеспечивающих воспроизводимость и сопоставимость результатов. Совместное применение метрик локализации, точности, полноты и надёжности идентификации формирует фундамент для объективного сравнения методов и их внедрения в прикладные системы компьютерного зрения.

Таблица 2

Основные метрики оценки систем детекции и распознавания лиц

Метрика	Подзадача	Описание
IoU	Детекция	Степень совпадения предсказанной и эталонной рамок
Precision	Детекция	Доля корректных обнаружений среди всех найденных
Recall	Детекция	Доля успешно найденных лиц относительно всех существующих
AP, mAP	Детекция	Средняя точность при разных порогах и уровнях сложности
TAR / FAR	Верификация (1:1)	Доля корректных и ложных приёмов при заданном пороге
EER	Верификация (1:1)	Точка равенства ложных приёмов и отказов, показатель надёжности
Rank-k / CMC	Идентификация (1:N)	Вероятность правильной идентификации в топ-k результатах
TPIR@FPIR	Идентификация (1:N)	Доля корректных идентификаций при фиксированной FPIR

Классические алгоритмы обработки изображений. Одной из ключевых задач в системах компьютерного зрения и биометрической идентификации является повышение качества изображений лиц, полученных в неблагоприятных условиях. Низкое освещение, шум, низкое разрешение и артефакты сжатия значительно осложняют распознавание и требуют применения эффективных методов предобработки. В научной и прикладной литературе особое внимание уделяется трём направлениям: глобальному выравниванию гистограммы (HE), его локализованной адаптации (CLAHE) и современным методам подавления шума на основе сверточных нейросетей (Denoising CNN). Рассмотрим принципы работы, преимущества и ограничения этих подходов, а также их возможную интеграцию в единую систему обработки.

Метод Histogram Equalization основан на перераспределении интенсивностей яркости изображения таким образом, чтобы гистограмма приблизилась к равномерному распределению. Это позволяет усилить контраст, особенно в темных и среднеярких областях, визуально улучшая изображение и выделяя ключевые элементы сцены. HE отличается простотой реализации и не требует параметров или обучающих данных, что делает его удобным для применения в системах реального времени при умеренных условиях съемки [15]. Однако у метода есть существенные недостатки. Глобальный характер обработки не позволяет различать области с разной значимостью, что может привести к пересвету ярких участков и усилению шума в темных. Метод не учитывает структуру изображения - например, форму глаз, рта или контуры лица - что ограничивает его эффективность в биометрических задачах. Кроме того, при низком качестве исходного изображения HE может усиливать артефакты, негативно влияя на точность последующей обработки [15].

Contrast Limited Adaptive Histogram Equalization (CLAHE) представляет собой улучшенную версию Adaptive Histogram Equalization (AHE), при которой изображение делится на мелкие регионы (тайлы), и в каждом из них строится локальная гистограмма, ограниченная по высоте (clip limit) [5]. Такой подход предотвращает чрезмерное усиление контраста и делает метод более устойчивым к шуму. Интерполяция между тайлами обеспечивает плавность переходов и сохраняет целостность изображения. CLAHE особенно эффективен в условиях неравномерного освещения, позволяя усилить локальные детали, которые часто теряются при глобальной обработке. Однако обработка производится по единому алгоритму для всех регионов изображения, включая фон, что может привести к усилению нерелевантных участков и маскировке значимых черт лица. Кроме того, эффективность CLAHE зависит от правильно подобранных параметров, при неверной настройке возможны как усиление шума, так и отсутствие заметного улучшения [16].

С появлением методов глубокого обучения стали развиваться обучаемые подходы к предобработке изображений. Denoising CNN (например, DnCNN, FFDNet) обучаются на парах "зашумлённое – чистое" изображение и способны эффективно восстанавливать текстуры, границы и детали даже при сложных шумовых искажениях. В отличие от классических фильтров, нейросети способны адаптироваться к профилю шума и сохранять важные структурные элементы лица. Одним из архитектурных преимуществ DnCNN является использование остаточного обучения: сеть предсказывает карту шума, которая затем вычитается из входного изображения. Это ускоряет обучение и повышает устойчивость к переобучению [17]. Однако использование нейросетевых методов сопряжено с высокой вычислительной нагрузкой: необходимы GPU и большие объёмы данных для обучения. Хотя предварительно обученные модели могут быть внедрены в реальные системы видеонаблюдения, разнообразие условий съёмки и сенсоров требует дополнительной валидации. Кроме того, при неудачной настройке модели возможно чрезмерное сглаживание, что может приводить к потере важных признаков [18].

Современные исследования направлены на создание end-to-end архитектур, объединяющих CLAFE, денойзинг и извлечение признаков в единую обучаемую структуру. Особое внимание уделяется включению attention-механизмов, многомасштабных свёрток и адаптивных стратегий подавления шума, учитывающих контекст сцены. Такие подходы позволяют минимизировать потери информации при сохранении вычислительной эффективности.

Исходя из изложенного выше, классические методы предобработки изображений, несмотря на простоту и низкие вычислительные затраты, требуют осторожного применения в задачах распознавания лиц. Их ограниченная адаптивность и отсутствие учёта семантики изображения могут снижать точность идентификации. Интеграция современных обучаемых решений значительно повышает устойчивость систем к внешним искажениям, обеспечивая более высокое качество входных данных для распознающих модулей.

Каскадные подходы. Одним из перспективных подходов для решения задачи улучшения изображений в условиях низкой освещённости является каскадная обработка, включающая этап генеративного усиления качества изображения и последующую идентификацию. В последние годы получили широкое распространение подходы, основанные на генеративных и обучаемых нейросетевых моделях. Одной из первых значимых разработок, ориентированных на такие задачи, стала LLNet, построенная на глубокой сверточной архитектуре автоэнкодера. Эта модель показала высокую эффективность в восстановлении деталей и усилении визуального качества изображений. Она обучается в режиме с учителем на синтетически затемнённых и зашумлённых изображениях, что позволяет ей корректировать как яркость, так и структуру сцены [19].

LLNet решает две задачи – коррекцию освещённости и подавление шумов - в рамках единой архитектуры. Такая комбинация позволяет модели адаптироваться к различным типам искажений без необходимости раздельного обучения отдельных компонентов. В процессе обучения LLNet минимизирует среднеквадратичную ошибку реконструкции между входным и целевым изображением. Использование глубоких слоёв и нелинейных активаций позволяет эффективно восстанавливать даже сильно деградированные изображения. Хотя LLNet обучается на ограниченных синтетических выборках, при достаточном разнообразии обучающих данных модель демонстрирует хорошие обобщающие способности, включая ранее не встречавшиеся шумы и уровни затемнения [19]. Однако в реальных задачах её эффективность может снижаться при отсутствии адаптации к новым условиям.

Другой важной моделью является EnlightenGAN, которая обучается в непарной постановке, не требуя точных пар «до/после». Это делает модель особенно гибкой и применимой в условиях, где невозможно собрать парные датасеты. Архитектура EnlightenGAN базируется на модифицированном U-Net с механизмами внимания и включает два дискриминатора: глобальный, контролирующий структуру сцены и освещённость в целом, и локальный, направленный на оценку мелких деталей, таких как текстура кожи и края объектов. Генератор выполняет задачу повышения яркости и контрастности, акцентируя внимание на ключевых областях сцены (например, лицо, глаза, контурные элементы).

Особенностью EnlightenGAN является сложная функция потерь, объединяющая традиционную состязательную компоненту с дополнительными термами, направленными на сохранение цветовой консистентности, структурной целостности и глобального стиля освещения [20]. Модель не требует дополнительных карт освещённости или ручных масок, что упрощает практическое внедрение. Несмотря на наличие сложной архитектуры, при оптимизации модели и предварительной подготовке inference может быть реализован в системах с ограниченными ресурсами, включая мобильные и встраиваемые устройства. Эффективность EnlightenGAN подтверждена на разнообразных датасетах (LOL, MIT-Adobe FiveK), включая сцены с различными уровнями освещённости [21]. Тем не менее, при использовании в задачах реального времени модель требует компромиссов между качеством и вычислительной нагрузкой, а улучшение визуального качества не всегда прямо коррелирует с точностью последующего распознавания.

Zero-DCE представляет собой модель, ориентированную на улучшение изображений с минимальной вычислительной нагрузкой и отсутствием зависимости от парных данных. В отличие от GAN-подходов, Zero-DCE не использует состязательное обучение, а реализует прямую регрессию кривых усиления освещённости (illumination adjustment curves), которые применяются к изображению на уровне пикселей [21]. Эти кривые предсказываются компактной свёрточной нейросетью и позволяют адаптивно корректировать яркость в различных областях изображения.

Обучение модели осуществляется с использованием многоцелевой функции потерь, включающей термы для сохранения цветовой консистентности, ограничения переэкспонирования, сглаженности кривых и предотвращения усиления артефактов. За счёт отказа от использования эталонных изображений модель легко адаптируется к новым данным и демонстрирует высокую скорость обработки. Это делает Zero-DCE особенно привлекательной для мобильных и встраиваемых систем. Однако отсутствие модуля подавления шумов ограничивает её применение в случаях с высокой зашумлённостью: усиление освещения может не только не устранить шум, но и усилить его визуальное проявление. Кроме того, на изображениях с яркими точечными источниками света (например, фары, уличные фонари) возможны локальные искажения. Таким образом, в задачах, где требуется одновременно подавление шума и коррекция освещённости, Zero-DCE рекомендуется использовать в сочетании с дополнительными методами предобработки.

Обобщая, все три рассмотренные модели демонстрируют разные подходы к решению проблемы улучшения изображений в условиях слабой освещённости. LLNet ориентирована на комбинированное решение задач восстановления яркости и подавления шумов, EnlightenGAN обеспечивает высокое качество за счёт генеративного моделирования, а Zero-DCE предлагает лёгкое и быстрое решение без необходимости в парных данных, при этом требует внешнего механизма подавления шумов для комплексной обработки.

RetinaFace – это одноэтапная модель обнаружения лиц, разработанная в рамках проекта InsightFace и отличающаяся высокой точностью и надёжностью при детекции лиц различных масштабов и в сложных условиях. Архитектура RetinaFace базируется на глубокой свёрточной сети, чаще всего ResNet-50, дополненной сетью пирамидальных признаков (Feature Pyramid Network, FPN), которая обеспечивает многомасштабную обработку входных изображений. Для повышения чувствительности к контексту и устойчивости к вариациям позы, в структуру FPN встроены контекстные модули, расширяющие рецептивное поле, а также деформируемые свёртки, адаптирующиеся к различным конфигурациям лицевых черт [22].

В отличие от многих моделей, RetinaFace выполняет не только локализацию лиц в виде ограничивающих рамок (bounding boxes), но и одновременно предсказывает координаты пяти ключевых точек: центров глаз, кончика носа и уголков рта. Эта особенность обеспечивает возможность точного выравнивания лиц, необходимого для последующих этапов идентификации и верификации. Для устойчивого обнаружения лиц разных размеров RetinaFace использует плотную сетку якорных рамок (anchor boxes), охватывающих широкий диапазон масштабов, что позволяет уверенно находить как крупные, так и мелкие лица, включая частично перекрытые и находящиеся в плохо освещённых или низко-контрастных областях изображения [23].

Модель обучается в многозадачной постановке, одновременно оптимизируя классификацию (наличие лица), регрессию ограничивающих рамок и локализацию ключевых точек. Такая стратегия повышает эффективность и точность работы модели в разнообразных сценариях. RetinaFace показывает превосходные результаты на сложных датасетах, таких как WIDER FACE, демонстрируя устойчивость к вариациям позы, освещения, наличию очков, масок и частичным окклюзиям. Благодаря объединению точной локализации и предсказания ключевых точек, RetinaFace является надёжной основой для подготовки изображений к системам верификации и распознавания.

После генеративного улучшения изображений (например, с использованием LLNet, EnlightenGAN или Zero-DCE), выровненные лица могут быть эффективно переданы на вход системам идентификации. Одними из наиболее результативных решений в этой области являются ArcFace и FaceNet. ArcFace реализует модифицированную функцию потерь на основе угловой дистанции (Additive Angular Margin Loss), которая усиливает межклассовую разницу между признаковыми векторами и тем самым повышает точность идентификации, особенно в условиях, близких к реальным. Такой подход делает модель менее чувствительной к внешним помехам, включая вариации выражения лица и освещённости.

FaceNet, в свою очередь, использует триплетную функцию потерь (triplet loss), направленную на минимизацию расстояния между представлениями одного и того же человека и максимизацию между различными. В результате FaceNet строит компактное и метрически устойчивое признаковое пространство, пригодное не только для идентификации, но и для задач кластеризации и поиска по базе. Обе модели требуют предварительного выравнивания лица по ключевым точкам, полученным от RetinaFace, что критически важно для достижения высокой точности в downstream-задачах.

Следовательно, RetinaFace не только обеспечивает точную локализацию лиц на изображениях, но и играет ключевую роль в подготовке входных данных для высокоточных систем идентификации, таких как ArcFace и FaceNet. Совместное использование этих моделей в едином пайплайне обеспечивает надёжную и устойчивую работу в реальных условиях съёмки.

На рис. 3 представлена схема формирования признаков в ArcFace. Изображение лица подаётся на вход backbone-сети, где извлекается нормализованный вектор признаков. Этот вектор подаётся на классификационный слой, формирующий финальное решение об идентичности (например, Иванов И.И.).

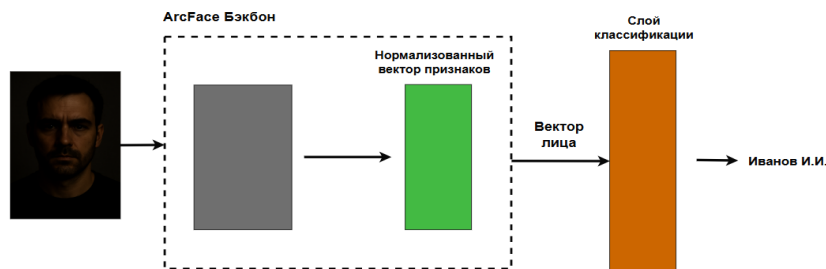


Рис. 3. Схема формирования признаков и классификации в ArcFace

Основное преимущество каскадного подхода заключается в его универсальности и способности значительно повысить точность распознавания лиц за счет предварительного улучшения изображений. Также данный подход обеспечивает адаптивность систем к разнообразным условиям эксплуатации без изменения базового алгоритма распознавания. Однако, несмотря на очевидные плюсы, каскадный подход обладает рядом недостатков. Основной из них – это возрастание вычислительной сложности и увеличение времени обработки изображений за счет дополнительных этапов улучшения качества. Кроме того, генеративные модели могут иногда вводить собственные артефакты, негативно влияющие на точность распознавания, особенно если модель выбрана неудачно или плохо обучена.

Принцип Identity Loss. Часто встречаемой проблемой при генерации или восстановлении изображений лиц остаётся сохранение индивидуальных черт человека, то есть его биометрической идентичности. Даже если модель успешно устраняет шум или корректирует освещение, результат может потерять сходство с исходным лицом. Для решения этой задачи в состав функции потерь современных генеративных сетей включают специальный компонент - identity loss, отвечающий за сохранение личности на уровне признаков. Его основная идея заключается в том, чтобы оценивать не просто сходство изображений по пикселям, а совпадение их описаний в пространстве высокоуровневых признаков, формируемых специализированными сетями распознавания лиц - например, ArcFace или FaceNet [22, 23]. То есть, модель учится изменять внешний вид изображения (яркость, контраст, экспозицию), не разрушая те особенности, которые делают лицо уникальным.

В отличие от традиционных метрик реконструкции, identity loss измеряет расстояние между индикативными векторами исходного и сгенерированного изображения. Чем меньше это расстояние, тем лучше генератор сохраняет идентичность. По сути, речь идёт о переносе критерия реалистичности с уровня пикселей на уровень семантических представлений. Такой подход позволяет системе быть менее чувствительной к изменению внешних условий съёмки и фокусироваться на постоянных биометрических характеристиках: пропорциях лица, конфигурации глаз и носа, характерных тенях и складках. Применение этого принципа необходимо при работе с низкоосвещёнными кадрами, где стандартные методы восстановления часто искажают форму или текстуру кожи.

Практическая реализация identity loss основана на совместном обучении генератора и заранее обученной модели-распознавателя, чьи параметры не изменяются в процессе обучения. Сравнивая эмбединги исходного и восстановленного изображений, система минимизирует их расхождение, тем самым заставляя генератор восстанавливать не просто схожее лицо, а то же самое. В некоторых работах вместо евклидовой меры используется косинусное сходство между векторами признаков, что позволяет точнее оценивать угловое расстояние в признаковом пространстве и повышает устойчивость модели к контрастным и цветовым искажениям.

Эффективность этого подхода подтверждена в ряде исследований. Так, в архитектуре FaceID-GAN [22] модуль идентификации используется наряду с генератором и дискриминатором и обеспечивает согласованность между визуальной реалистичностью и сохранением личности. Модель обучается на парах «искажённое – эталонное» изображение и демонстрирует значительное повышение точности распознавания после генерации. Похожие идеи реализованы и в LightenFace [24], где идентификационная потеря используется совместно с перцепционной и состязательной. Это позволяет улучшать освещённость лица без потери узнаваемости и при этом избегать чрезмерного сглаживания деталей. Использование идентификационного компонента в функции потерь помогает избежать «размывания» текстур и повышает метрики структурного сходства (SSIM) и восприятия (LPIPS), что напрямую связано с улучшением качества распознавания после предварительной обработки [20, 24].

Принцип identity loss служит мостом между задачами визуального улучшения и задачами идентификации. Он делает возможным создание генеративных систем, которые не только улучшают изображение, но и сохраняют саму личность человека, изображённого на ней. В контексте распознавания лиц при низкой освещённости это имеет особое значение: модель не просто корректирует яркость, а восстанавливает лицо так, чтобы его признаки остались согласованными с исходным представлением в пространстве ArcFace. Благодаря этому подходу удаётся добиться сочетания визуального качества и биометрической достоверности, что открывает путь к построению гибридных архитектур, объединяющих улучшение, детекцию и распознавание в единую согласованную систему [22, 23].

Примеры моделей с контролем идентичности. FaceID-GAN – это генеративная модель, предназначенная для восстановления изображений лиц с сохранением идентичности. Архитектура объединяет генератор, дискриминатор, модуль идентификации, attention-механизмы и адаптивный блок шумоподавления. Генератор построен по схеме U-Net и восстанавливает изображение из зашумлённого входа [24]. Attention-механизмы фокусируют обработку на ключевых зонах лица. Адаптивный шумоподавляющий блок

минимизирует шум, не затрагивая биометрически значимые признаки. Дискриминатор оценивает реалистичность сгенерированных изображений. Ключевым элементом является модуль идентификации, основанный на ArcFace [24]. Он извлекает признаки из оригинального и сгенерированного изображения и вычисляет идентификационную потерю как евклидово расстояние между векторами признаков. Это обеспечивает контроль идентичности на этапе генерации.

$$L_{\text{total}} = \lambda_{\text{adv}} L_{\text{adv}} + \lambda_{\text{id}} L_{\text{id}} + \lambda_{\text{perc}} L_{\text{perc}} + \lambda_{\text{den}} L_{\text{den}} + \lambda_{\text{att}} L_{\text{att}} \quad (12)$$

L_{adv} – соревновательная потеря; L_{id} – идентификационная; L_{perc} – перцепционная; L_{den} – шумоподавляющая; L_{att} – потеря внимания.

Веса λ настраиваются эмпирически.

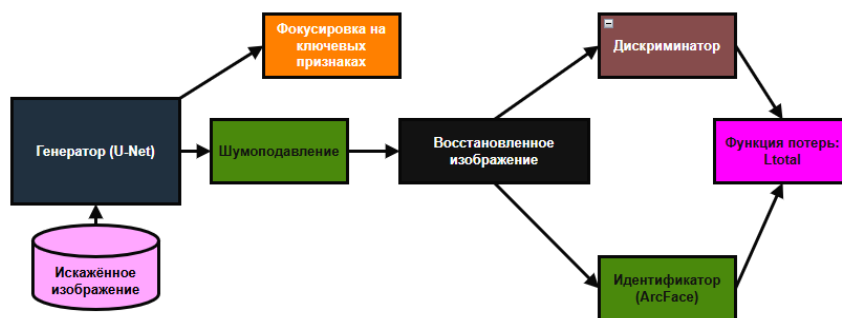


Рис. 4. Архитектура FaceID-GAN с модулями контроля идентичности и шумоподавления

Модель обучается end-to-end. На вход подаётся искажённое изображение, генератор восстанавливает лицо, дискриминатор оценивает реалистичность, а модуль идентификации – сохранение личности. Ошибки суммируются в итоговую функцию потерь.

В основе KinD++ лежит концепция разложения изображения на компоненты освещенности и отражения, что позволяет изолировать влияние освещения от текстурных особенностей сцены [25]. Модель обучается корректировать компонент освещенности, устраняя неравномерности и недостаточную яркость, в то время как компонент отражения сохраняет важные структурные детали. Такой подход обеспечивает целенаправленное усиление яркости без разрушения мелких текстур. KinD++ использует многоуровневую архитектуру, включающую три ключевых блока: модуль разложения, модуль восстановления освещенности и блок финальной доработки. Сначала входное изображение обрабатывается сетью разложения, которая с помощью сверточных слоев выделяет карты отражения и освещенности [26, 27]. Далее коррекция освещения проводится в пространстве яркости (luminance), полученной из RGB-изображения, что позволяет сохранить текстурные детали без искажений.

На завершающем этапе происходит объединение полученных компонентов с учетом локальных особенностей сцены – это помогает устранить оставшиеся артефакты и улучшить общее визуальное восприятие. Важной задачей модели является сохранение идентичности объектов, особенно в области лиц и сложных текстур. Для этого в функцию потерь включены perceptual loss, отслеживающий различия на уровне высокоуровневых признаков, и SSIM – показатель структурного сходства, поддерживающий локальную целостность изображения [28].

Дополнительно используется refinement-сеть – она «дошлифовывает» изображение, устраняя остаточные шумы и повышая резкость в важных зонах. Этот блок компенсирует возможные недочеты предыдущих этапов и делает результат визуально естественным. Для повышения реалистичности и стабильности обучения в KinD++ также применяются методы генеративно-состязательного обучения (GAN), где дискриминатор учится отличать реальные изображения от сгенерированных, заставляя генератор создавать более правдоподобные текстуры и устранять синтетические дефекты [17, 28]. Обучение модели

осуществляется на специально подготовленных датасетах с разнообразными условиями освещенности, что обеспечивает высокую обобщающую способность и устойчивость к различным сценариям применения. Благодаря совмещению методов физически обоснованного разложения изображения, идентичностно-ориентированных потерь и глубокой архитектуры с модулями доработки, KinD++ эффективно решает задачу восстановления изображений в условиях низкой освещенности.

Модель LightenFace предназначена для решения проблемы недостаточной освещенности изображений лиц. Её основная особенность заключается в объединении задач улучшения изображения и извлечения биометрических признаков в единую архитектуру [29]. В отличие от каскадных подходов, где сначала выполняется усиление изображения с помощью отдельной модели, а затем распознавание лица, LightenFace реализует совместное обучение, что позволяет устранить накопление ошибок, характерное для каскадных методов, и улучшить обобщающую способность модели [30]. Она особенно хорошо работает в условиях, где требуется высокая точность при ограниченных вычислительных возможностях [11]. Архитектура модели основана на принципе многозадачного обучения и сочетает в себе два основных модуля: один отвечает за улучшение качества изображения, другой – за извлечение признаков лица. Модуль улучшения устраняет типичные проблемы, возникающие при слабом освещении, такие как пониженный контраст, искажения яркости и шумы, подготавливая изображение для более точного анализа. После предварительного улучшения изображение передаётся во второй модуль, который извлекает устойчивые к освещению признаки лица, пригодные для идентификации. Обучение всей модели осуществляется с применением совокупной функции потерь, в которую входят компоненты, отвечающие за качество реконструкции изображения, перцептивное соответствие улучшенного изображения оригиналу, а также идентификационную точность на уровне векторных представлений. Модель демонстрирует высокие результаты на специализированных датасетах, таких как DarkFace, LFW-low-light и SCface. Благодаря тому, что LightenFace использует совместное обучение, она демонстрирует устойчивость к шумам и искажениям, характерным для условий ночного видеонаблюдения и съёмки с недостаточной экспозицией. При этом архитектура остаётся достаточно лёгкой для работы на мобильных и встраиваемых устройствах. Размер модели в сжатом виде может занимать менее 10 мегабайт, а сами вычисления могут быть выполнены в формате FP16, что критически важно для edge-вычислений [11, 25]. Технически модель реализует модифицированные свёрточные блоки с низкой вычислительной сложностью, использует методы нормализации с возможностью восстановления исходных характеристик цвета и обучается как на синтетически затемнённых, так и на реальных изображениях. Это позволяет ей адаптироваться к различным источникам данных без необходимости в дополнительной калибровке или ручной настройке параметров.

Identity loss, используемый в генеративных моделях (FaceID-GAN, EnlightenGAN), доказал эффективность в сохранении идентичности при аугментации данных или коррекции освещения. Однако его применение ограничено в условиях низкого качества изображений: шум и нечёткость признаков затрудняют извлечение идентификационных характеристик, а вычислительная сложность интеграции дополнительных сетей для расчёта потерь осложняет использование в реальном времени.

Механизмы внимания (spatial, channel, multi-scale) позволяют моделям фокусироваться на ключевых областях лица (глаза, нос, рот), минимизируя влияние фоновых искажений. Интеграция модулей типа SE-Block и CBAM усиливает значимость релевантных признаков, а self-attention механизмы, заимствованные из архитектур трансформеров, обеспечивают глобальный анализ контекста, повышая устойчивость к локальным искажениям. Дополнительный эффект достигается за счёт адаптивного шумоподавления: attention-guided denoising сочетает карты внимания с селективной фильтрацией, подавляя шум в нерелевантных зонах, а многоступенчатые автоэнкодеры с обратной связью динамически адаптируются к типу шума [27, 29].

Несмотря на прогресс, сохраняются компромиссы между качеством и ресурсоёмкостью. Генеративные модели (Zero-DCE, LLNet) эффективно улучшают освещение, но требуют значительных вычислений, а методы на основе CNN с многоуровневым вниманием, хотя и менее затратны, могут терять детализацию в экстремальных условиях [26, 28]. Оптимальным решением остаётся комбинация подходов: предобработка изображений с использованием lightweight-архитектур, адаптивная фильтрация и акцент на сохранение идентичности через модифицированные функции потерь.

Существующие подходы не могут в полной мере решать задачу распознавания лиц в условиях низкой освещённости, где важна высокая точность и сохранение биометрической идентичности. Для этого необходима интеграция нескольких технологий в единую систему с возможностью совместной оптимизации. В связи с этим, предлагаемая концепция гибридной end-to-end архитектуры направлена на решение этих проблем с использованием инновационных методов.

Концепция предлагаемой гибридной end-to-end архитектуры. Предлагаемый проект представляет собой гибридную end-to-end архитектуру, предназначенную для эффективного распознавания лиц в условиях низкой освещённости. Концептуальной основой является интеграция нескольких специализированных модулей обработки и анализа изображений в едином графе вычислений. В состав системы входят генеративная модель EnlightenGAN, предназначенная для улучшения освещённости и подавления шумов, механизм идентификации ArcFace для точного распознавания личности и детектор RetinaFace для надёжного определения положения лиц на изображении.

Особенностью предлагаемой концепции является совместное обучение всех модулей как единого целого, что позволяет минимизировать накопление ошибок на каждом этапе обработки. Использование адаптивного шумоподавления и attention-механизмов способствует дополнительной адаптации системы к условиям съёмки, акцентируя внимание модели на наиболее важных участках лица (глаза, нос, рот).

Архитектура функционирует следующим образом: исходное изображение с низкой освещённостью поступает в генератор EnlightenGAN, который улучшает его качество и устраняет шумы. После предварительной обработки улучшенное изображение передаётся в модуль распознавания ArcFace, извлекающий уникальные признаки лица и идентифицирующий личность. Одновременно детектор RetinaFace определяет положение лица на изображении и оценивает достоверность обнаружения. Совместная оптимизация всех компонентов достигается с помощью комплексной функции потерь, включающей компоненты для учёта идентичности, детекции, а также визуальной реалистичности изображения.

В предложенной архитектуре, низкоосвещённый кадр I_{dark} сначала обрабатывается генератором EnlightenGAN, внутри которого выделены два ключевых под-блока: адаптивное шумоподавление и face-attention. На выходе формируется улучшенное изображение I_{enh} , поступающее одновременно в детектор RetinaFace и распознаватель ArcFace. Каждая ветвь генерирует специализированную функцию потерь (Detection Loss, Identity Loss и др.).

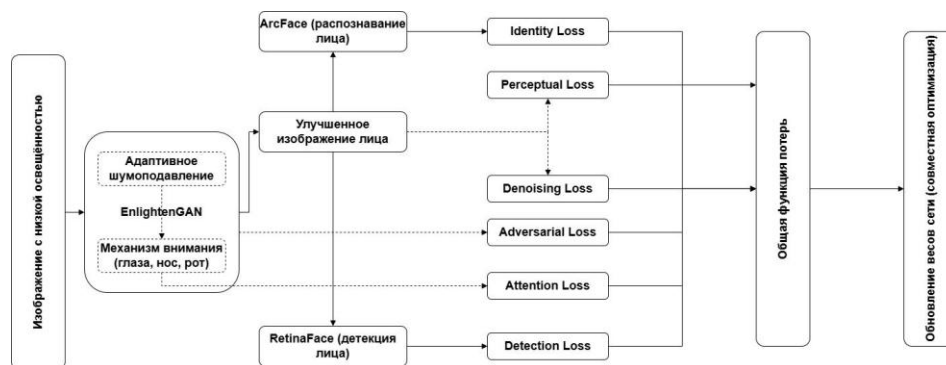


Рис. 5. Совместный пайплайн улучшения и анализа лицевых изображений в условиях низкой освещённости

Все потери агрегируются в общую функцию L_{total} , после чего единым градиентом обновляются веса всей сети. Совокупная функция потерь записывается как:

$$L_{\text{total}} = \lambda_{\text{adv}} L_{\text{adv}} + \lambda_{\text{perc}} L_{\text{perc}} + \lambda_{\text{den}} L_{\text{den}} + \lambda_{\text{id}} L_{\text{id}} + \lambda_{\text{det}} L_{\text{det}} + \lambda_{\text{att}} L_{\text{att}} \quad (13)$$

L_{adv} – Adversarial Loss (адверсариальная потеря): отвечает за реалистичность генерируемых изображений, стимулируя генератор к созданию изображений, неотличимых от реальных. Весовой коэффициент λ_{adv} определяет значимость этой потери.

L_{perc} – Perceptual Loss (перцепционная потеря): измеряет различие между входным и улучшенным изображением на уровне признаков, извлечённых из предобученной сверточной нейросети. Обеспечивает сохранение структуры и содержания изображения. λ_{perc} регулирует её влияние.

L_{den} – Denoising Loss (потеря подавления шума): отвечает за устранение шумов и артефактов в изображении, особенно критичных при повышении освещённости. λ_{den} задаёт силу регуляризации по шуму.

L_{id} – Identity Loss (потеря идентичности): оценивает, насколько хорошо сохранена личность на изображении после улучшения, сравнивая признаки, извлечённые ArcFace, с эталоном. λ_{id} отвечает за сохранение биометрической достоверности.

L_{det} – Detection Loss (потеря детекции): рассчитывается на основе выхода RetinaFace и отражает корректность обнаружения лица на улучшенном изображении. λ_{det} определяет значимость точной локализации лица.

L_{att} – Attention Loss (потеря внимания): побуждает генератор акцентировать внимание на ключевых чертах лица, таких как глаза, нос и рот. λ_{att} контролирует приоритет фокусировки на этих областях.

Такой трёхконтурный граф обеспечивает совместную оптимизацию задач повышения качества, детекции и идентификации лица, что критически важно для ночных и слабоосвещённых сцен.

Предлагаемая концепция характеризуется высоким инновационным потенциалом, поскольку аналогичные интегрированные системы на текущий момент практически не представлены в открытых источниках. Перспективы проекта заключаются в повышении точности и надёжности распознавания лиц при слабом освещении, что позволит значительно расширить области применения данной технологии, включая системы видеонаблюдения, контроль доступа и обеспечение безопасности в различных условиях.

Заключение. В обзоре проанализированы ключевые метрики для оценки качества улучшения изображения и точности распознавания, современные подходы к распознаванию лиц в условиях низкой освещённости. Традиционные методы, такие как Histogram Equalization (HE) и Contrast Limited Adaptive Histogram Equalization (CLAHE), демонстрируют высокую вычислительную эффективность, но недостаточно сохраняют идентичность лиц и устойчивы к шумам. В отличие от них, глубокие нейронные сети – Denoising CNN, LLNet, EnlightenGAN и Zero-DCE – обеспечивают значительное улучшение качества изображений, повышая точность последующей идентификации.

Особое внимание уделяется моделям с применением identity loss (FaceID-GAN, KinD++, LightenFace), которые сохраняют уникальные черты лиц, что критично для биометрических систем.

Одним из самых перспективных направлений остаётся создание гибридной end-to-end архитектуры, которая бы совмещала генеративные методы, механизмы внимания и адаптивное шумоподавление. В дальнейших исследованиях планируется реализация этой модели, её тестирование на расширенных датасетах, а также адаптация под работу в режиме реального времени. Отдельное внимание будет уделено снижению вычислительной нагрузки и настройке системы для работы на устройствах с ограниченными ресурсами – таких как мобильные платформы и встраиваемые системы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Лопатин Д.В.* Исследование эффективности дескрипторов особых точек на различных типах изображений: магистерская диссертация. – Тольятти: ТГУ, 2022. – 75 с.
2. *Гончаров А.В., Каркищенко А.Н.* Влияние освещённости на качество распознавания фронтальных лиц // Известия ЮФУ. Технические науки. – 2008. – № 4 (81). – С. 88-92.
3. *Liang J., Li Y., Yang M., Xu Y.* Recurrent Exposure Generation for Low-Light Face Detection // arXiv. – 2020. – arXiv:2007.10963. – doi: 10.48550/arXiv.2007.10963.
4. *Ying Z., Li G., Gao W.* A New Low-Light Image Enhancement Algorithm Using Camera Response Model // Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV Workshops). – 2017. – P. 3015-3023. – doi: 10.1109/ICCVW.2017.354.

5. Wei C., Wang W., Yang W., Liu J. Deep Retinex Decomposition for Low-Light Enhancement (Retinex-Net) // arXiv. – 2018. – arXiv:1808.04560. – doi: 10.48550/arXiv.1808.04560.
6. Chen C., Chen Q., Xu J., Koltun V. Learning to See in the Dark (SID) // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2018. – P. 3291-3300. – doi: 10.1109/CVPR.2018.00347.
7. Jiang Y., Gong X., Liu D., Cheng Y., Fang C., Shen X., Yang J. Low-Light Image and Video Enhancement: A Comprehensive Survey // arXiv. – 2022. – arXiv:2212.10772. – doi: 10.48550/arXiv.2212.10772.
8. Singh R., Kapoor K. Contrast Enhancement Techniques: A Survey // International Journal of Signal and Image Processing (IJSIP). – 2017. – Vol. 8, No. 3. – P. 45-56.
9. Kundu D., Chaudhuri B. Enhancement Measure Estimation (EME): A Metric for Image Contrast Evaluation // IET Image Processing. – 2014. – Vol. 8, No. 11. – P. 731-739. – doi: 10.1049/iet-ipr.2013.0546.
10. Ковалевский В.А. Методы улучшения изображений в системах видеонаблюдения при недостаточной освещённости: диссертация. – М.: МГТУ им. Баумана, 2021. – 112 с.
11. 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 (Zero-DCE++) // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2022. – Vol. 44, No. 11. – P. 8825-8838. – doi: 10.1109/TPAMI.2022.3141833.
12. Yang S., Luo P., Loy C. C., Tang X. WIDER FACE: A Face Detection Benchmark // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – P. 5525-5533. – doi: 10.1109/CVPR.2016.596.
13. Jain V., Learned-Miller E. FDDB: A Benchmark for Face Detection in Unconstrained Settings // University of Massachusetts Amherst Technical Report. – 2010. – 32 p.
14. Maze B., Adams J., Duncan J. A., Kalka N., Grother P., Ngan M., Patterson E., Kao H.-C. IARPA Janus Benchmark-C: Face Dataset and Protocol // Proceedings of the International Conference on Biometrics (ICB). – 2018. – P. 1-8. – doi: 10.1109/ICB2018.2018.00012.
15. Patel O., Maravi Y.P.S., Sharma S. A Comparative Study of Histogram Equalization Based Image Enhancement Techniques for Brightness Preservation and Contrast Enhancement // Signal & Image Processing: An International Journal. – 2013. – Vol. 4, No. 1. – P. 1-10.
16. Bulat A., Tzimiropoulos G. How Far Are We from Solving the 2D & 3D Face Alignment Problem? // Proceedings of the IEEE International Conference on Computer Vision (ICCV). – 2017. – P. 2488-2497. – doi: 10.1109/ICCV.2017.270.
17. Yamashita R., Nishio M., Do R.K.G., Togashi K. Convolutional Neural Networks: An Overview and Application in Radiology // Insights into Imaging. – 2018. – Vol. 9. – P. 611-629. – doi: 10.1007/s13244-018-0639-9.
18. Lore K.G., Akintayo A., Sarkar S. LLNet: A Deep Autoencoder Approach to Natural Low-Light Image Enhancement // Proceedings of the IEEE International Conference on Image Processing (ICIP). – 2017. – P. 1422-1426. – doi: 10.1109/ICIP.2017.8296512.
19. Jiang Y., Gong X., Liu D., Cheng Y., Fang C., Shen X., Yang J. EnlightenGAN: Deep Light Enhancement without Paired Supervision // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2021. – P. 2862-2871. – doi: 10.1109/CVPR46437.2021.00473.
20. Deng J., Guo J., Zhou Y., Yu J., Kotsia I., Zafeiriou S. RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2020. – P. 5203-5212. – doi: 10.1109/CVPR42600.2020.00525.
21. Shen Y., Liu H., Qi G.-J., Song J., Tao D. FaceID-GAN: Learning a Symmetry Three-Player GAN for Identity-Preserving Face Synthesis // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2018. – P. 821-830. – doi: 10.1109/CVPR.2018.00092.
22. Li C., Guo C., Loy C.C., Cong R. LightenFace: Realistic Face Relighting via GAN with Identity Loss // IEEE Transactions on Image Processing. – 2022. – Vol. 31. – P. 2551-2564. – doi: 10.1109/TIP.2022.3157457.
23. Zamir S.W., Arora A., Khan S., Hayat M., Khan F.S., Yang M.H., Shao L. Learning Enriched Features for Real Image Restoration and Enhancement // European Conference on Computer Vision (ECCV). – 2020. – P. 492-511. – doi: 10.1007/978-3-030-58595-2_30.
24. Shi Y., Wang Z., Tang Z., Wang J., Yin Y. LighterFace Model for Community Face Detection and Recognition // IEEE Access. – 2021. – Vol. 9. – P. 50752-50762. – doi: 10.1109/ACCESS.2021.3069343.
25. Singh P., Reibman A.R. Task-Aware Image Quality Estimators for Face Detection // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 6211-6220. – doi: 10.1109/CVPR52688.2022.00614.
26. Liang X., Chen X., Ren K., Miao X., Chen Z., Jin Y. Low-Light Image Enhancement via Adaptive Frequency Decomposition Network // Scientific Reports. – 2023. – Vol. 13. – Article No. 15208. – doi: 10.1038/s41598-023-40899-8.

27. Zhang Y., Zhang J., Guo X. Kindling the Darkness: A Practical Low-Light Image Enhancer // *Proceedings of the International Conference on Multimedia Modeling (MMM)*. – 2021. – P. 133-145. – doi: 10.1007/978-3-030-67832-6_12.
28. Xie Y., Ou J., Wen B., Yu Z., Tian W. A Joint Learning Method for Low-Light Facial Expression Recognition // *Complex & Intelligent Systems*. – 2024. – Vol. 10. – P. 2595-2611. – doi: 10.1007/s40747-024-01762-z.
29. Grother P., Ngan M., Hanaoka K. Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects // *NIST Interagency Report 8280*. – 2019. – URL: <https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf> (дата обращения: 23.04.2026).
30. George A., Ecabert C., Shahreza H.O., Kotwal K., Marcel S. EdgeFace: Efficient Face Recognition Model for Edge Devices // *arXiv*. – 2023. – arXiv:2307.01838. – doi: 10.48550/arXiv.2307.01838.

REFERENCES

1. Lopatin D.V. Issledovanie effektivnosti deskriptorov osobykh toчек na razlichnykh tipakh izobrazheniy: masterskaya dissertatsiya [Study of the Efficiency of Feature Point Descriptors on Various Types of Images: Master's thesis]. Tol'yatti: TGU, 2022, 75 p.
2. Goncharov A.V., Karkishchenko A.N. Vliyanie osveshchennosti na kachestvo raspoznavaniya frontal'nykh lits [Influence of illumination on the quality of frontal face recognition], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2008, No. 4 (81), pp. 88-92.
3. Liang J., Li Y., Yang M., Xu Y. Recurrent Exposure Generation for Low-Light Face Detection, *arXiv*, 2020. arXiv:2007.10963. doi: 10.48550/arXiv.2007.10963.
4. Ying Z., Li G., Gao W. A New Low-Light Image Enhancement Algorithm Using Camera Response Model, *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, 2017, pp. 3015-3023. doi: 10.1109/ICCVW.2017.354.
5. Wei C., Wang W., Yang W., Liu J. Deep Retinex Decomposition for Low-Light Enhancement (Retinex-Net), *arXiv*, 2018. arXiv:1808.04560. doi: 10.48550/arXiv.1808.04560.
6. Chen C., Chen Q., Xu J., Koltun V. Learning to See in the Dark (SID), *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 3291-3300. doi: 10.1109/CVPR.2018.00347.
7. Jiang Y., Gong X., Liu D., Cheng Y., Fang C., Shen X., Yang J. Low-Light Image and Video Enhancement: A Comprehensive Survey, *arXiv*, 2022. arXiv:2212.10772. doi: 10.48550/arXiv.2212.10772.
8. Singh R., Kapoor K. Contrast Enhancement Techniques: A Survey, *International Journal of Signal and Image Processing (IJSIP)*, 2017, Vol. 8, No. 3, pp. 45-56.
9. Kundu D., Chaudhuri B. Enhancement Measure Estimation (EME): A Metric for Image Contrast Evaluation, *IET Image Processing*, 2014, Vol. 8, No. 11, pp. 731-739. doi: 10.1049/iet-ipr.2013.0546.
10. Kovalevskiy V.A. Metody uluchsheniya izobrazheniy v sistemakh videonablyudeniya pri nedostatatochnoy osveshchennosti: dissertatsiya [Methods for improving images in video surveillance systems under low illumination conditions: dissertation]. Moscow: MGTU im. Baumana, 2021, 112 p.
11. 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 (Zero-DCE++), *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, Vol. 44, No. 11, pp. 8825-8838. doi: 10.1109/TPAMI.2022.3141833.
12. Yang S., Luo P., Loy C. C., Tang X. WIDER FACE: A Face Detection Benchmark, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 5525-5533. doi: 10.1109/CVPR.2016.596.
13. Jain V., Learned-Miller E. FDDB: A Benchmark for Face Detection in Unconstrained Settings, *University of Massachusetts Amherst Technical Report*, 2010, 32 p.
14. Maze B., Adams J., Duncan J. A., Kalka N., Grother P., Ngan M., Patterson E., Kao H.-C. IARPA Janus Benchmark-C: Face Dataset and Protocol, *Proceedings of the International Conference on Biometrics (ICB)*, 2018, pp. 1-8. doi: 10.1109/ICB2018.2018.00012.
15. Patel O., Maravi Y.P.S., Sharma S. A Comparative Study of Histogram Equalization Based Image Enhancement Techniques for Brightness Preservation and Contrast Enhancement, *Signal & Image Processing: An International Journal*, 2013, Vol. 4, No. 1, pp. 1-10.
16. Bulat A., Tzimiropoulos G. How Far Are We from Solving the 2D & 3D Face Alignment Problem?, *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2488-2497. doi: 10.1109/ICCV.2017.270.
17. Yamashita R., Nishio M., Do R.K.G., Togashi K. Convolutional Neural Networks: An Overview and Application in Radiology, *Insights into Imaging*, 2018, Vol. 9, pp. 611-629. doi: 10.1007/s13244-018-0639-9.

18. Lore K.G., Akintayo A., Sarkar S. LLNet: A Deep Autoencoder Approach to Natural Low-Light Image Enhancement, *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 1422-1426. doi: 10.1109/ICIP.2017.8296512.
19. Jiang Y., Gong X., Liu D., Cheng Y., Fang C., Shen X., Yang J. EnlightenGAN: Deep Light Enhancement without Paired Supervision, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 2862-2871. doi: 10.1109/CVPR46437.2021.00473.
20. Deng J., Guo J., Zhou Y., Yu J., Kotsia I., Zafeiriou S. RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 5203-5212. doi: 10.1109/CVPR42600.2020.00525.
21. Shen Y., Liu H., Qi G.-J., Song J., Tao D. FaceID-GAN: Learning a Symmetry Three-Player GAN for Identity-Preserving Face Synthesis, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 821-830. doi: 10.1109/CVPR.2018.00092.
22. Li C., Guo C., Loy C.C., Cong R. LightenFace: Realistic Face Relighting via GAN with Identity Loss, *IEEE Transactions on Image Processing*, 2022, Vol. 31, pp. 2551-2564. doi: 10.1109/TIP.2022.3157457.
23. Zamir S.W., Arora A., Khan S., Hayat M., Khan F.S., Yang M.H., Shao L. Learning Enriched Features for Real Image Restoration and Enhancement, *European Conference on Computer Vision (ECCV)*, 2020, pp. 492-511. doi: 10.1007/978-3-030-58595-2_30.
24. Shi Y., Wang Z., Tang Z., Wang J., Yin Y. LighterFace Model for Community Face Detection and Recognition, *IEEE Access*, 2021, Vol. 9, pp. 50752-50762. doi: 10.1109/ACCESS.2021.3069343.
25. Singh P., Reibman A.R. Task-Aware Image Quality Estimators for Face Detection, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 6211-6220. doi: 10.1109/CVPR52688.2022.00614.
26. Liang X., Chen X., Ren K., Miao X., Chen Z., Jin Y. Low-Light Image Enhancement via Adaptive Frequency Decomposition Network, *Scientific Reports*, 2023, Vol. 13, Article No. 15208. doi: 10.1038/s41598-023-40899-8.
27. Zhang Y., Zhang J., Guo X. Kindling the Darkness: A Practical Low-Light Image Enhancer, *Proceedings of the International Conference on Multimedia Modeling (MMM)*, 2021, pp. 133-145. doi: 10.1007/978-3-030-67832-6_12.
28. Xie Y., Ou J., Wen B., Yu Z., Tian W. A Joint Learning Method for Low-Light Facial Expression Recognition, *Complex & Intelligent Systems*, 2024, Vol. 10, pp. 2595-2611. doi: 10.1007/s40747-024-01762-z.
29. Grother P., Ngan M., Hanaoka K. Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects, *NIST Interagency Report 8280*, 2019. Available at: <https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf> (accessed 23 April 2026).
30. George A., Ecabert C., Shahreza H.O., Kotwal K., Marcel S. EdgeFace: Efficient Face Recognition Model for Edge Devices, *arXiv*, 2023. arXiv:2307.01838. doi: 10.48550/arXiv.2307.01838.

Морозов Дмитрий Александрович – Волгоградский государственный технический университет; e-mail: dimka5rus@yandex.ru; г. Волгоград, Россия; тел.: +79823643521; кафедра программного обеспечения автоматизированных систем; студент.

Гилка Вадим Викторович – Волгоградский государственный технический университет; e-mail: gilka_vv@mail.ru; г. Волгоград, Россия; тел.: +79996247154; кафедра программного обеспечения автоматизированных систем; к.т.н.; старший преподаватель.

Кузнецова Агнесса Сергеевна – Волгоградский государственный технический университет; e-mail: agnessakyz@yandex.ru; г. Волгоград, Россия; тел.: +79275174399; кафедра программного обеспечения автоматизированных систем; старший преподаватель.

Morozov Dmitry Alexandrovich – Volgograd State Technical University; e-mail: dimka5rus@yandex.ru; Volgograd, Russia; phone: +79823643521; the Department of Software for Automated Systems; student.

Gilka Vadim Viktorovich – Volgograd State Technical University; e-mail: gilka_vv@mail.ru; Volgograd, Russia; phone: +79996247154; the Department of Software for Automated Systems; cand. of eng. sc.; senior lecturer.

Kuznetsova Agnessa Sergeevna – Volgograd State Technical University; e-mail: agnessakyz@yandex.ru; Volgograd, Russia; phone: +79275174399; the Department of Software for Automated Systems; senior lecturer.