

**ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ/INFORMATICS AND INFORMATION PROCESSES**DOI: <https://doi.org/10.60797/IRJ.2026.168.4> EDN: QNELVC**АРХИТЕКТУРНОЕ ОБОСНОВАНИЕ СКВОЗНОЙ МОДЕЛИ ДЛЯ ДЕТЕКЦИИ И РАСПОЗНАВАНИЯ ЛИЦ В УСЛОВИЯХ НИЗКОЙ ОСВЕЩЁННОСТИ С СОХРАНЕНИЕМ ИДЕНТИЧНОСТИ**

Научная статья

Морозов Д.А.¹, Гилка В.В.², Кузнецова А.С.^{3,*}, Громов Е.Г.⁴²ORCID : 0000-0003-1470-8766;³ORCID : 0000-0001-5554-3887;^{1,2,3,4} Волгоградский государственный технический университет, Волгоград, Российская Федерация

* Корреспондирующий автор (agnessakyz[at]yandex.ru)

Предложена: 20.03.2026; Принята: 26.05.2026; Опубликовано: 17.06.2026

Аннотация

В работе представлено архитектурное решение задачи сквозной детекции и распознавания лиц, функционирующее в условиях недостаточной освещенности. Низкий уровень освещения вызывает критическое ухудшение качества входных данных: рост шумов, падение контраста и потерю текстурной информации, что негативно влияет на точность локализации объектов и стабильность извлекаемых биометрических эмбедингов. Для преодоления указанных ограничений предлагается интегрировать модуль улучшения изображения, детектор и распознаватель в единый вычислительный контур. Такая структура, в отличие от классических каскадных решений с внешней предобработкой, обеспечивает согласованное обучение всех компонентов и целенаправленную оптимизацию признакового пространства, релевантного для сохранения identity-информации.

Центральная методологическая проблема при реализации данного подхода заключается в некорректности использования нестабильного эмбединга из зашумленного темного кадра в качестве эталона для функции потерь идентичности. В работе анализируются два способа разрешения этого противоречия. Первый предполагает парное обучение с применением эталонных изображений нормальной освещенности. Второй, не требующий наличия светлых эталонов, базируется на механизме teacher–student дистилляции с momentum teacher, что позволяет стабилизировать процесс формирования биометрических представлений. Дополнительно рассматриваются аспекты стабилизации ключевых точек, гейтинг потерь на основе уверенности детектора и балансировка многокомпонентной целевой функции. Предложенная архитектура создает методологически обоснованную основу для разработки программного комплекса, ориентированного на высокую точность распознавания личности в условиях сложных световых сцен.

Ключевые слова: низкая освещённость, детекция лиц, распознавание лиц, сохранение идентичности, общий энкодер, teacher student дистилляция, гейтинг потерь.

ARCHITECTURAL SUBSTANTIATION OF AN END-TO-END MODEL FOR FACE DETECTION AND RECOGNITION IN LOW-LIGHT CONDITIONS WHILE PRESERVING IDENTITY

Research article

Morozov D.A.¹, Gilka V.V.², Kuznecova A.S.^{3,*}, Gromov Y.G.⁴²ORCID : 0000-0003-1470-8766;³ORCID : 0000-0001-5554-3887;^{1,2,3,4} Volgograd State Technical University, Volgograd, Russian Federation

* Corresponding author (agnessakyz[at]yandex.ru)

Suggested: 20.03.2026; Accepted: 26.05.2026; Published: 17.06.2026

Abstract

The work presents an architectural solution for end-to-end face detection and recognition that operates in low-light conditions. Low light levels cause a critical deterioration in the quality of the input data: increased noise, reduced contrast and loss of textural information, which negatively affects the accuracy of object localisation and the stability of the extracted biometric embeddings. To overcome these limitations, we suggest integrating an image enhancement module, a detector and a recogniser into a single computational loop. Unlike traditional cascaded solutions with external pre-processing, this structure ensures coordinated training of all components and targeted optimisation of the feature space relevant for preserving identity information.

The central methodological issue in implementing this approach lies in the incorrect use of unstable embedding from a noisy dark frame as a reference for the identity loss function. The paper analyses two methods for resolving this contradiction. The first involves paired learning using reference images with normal illumination. The second, which does not require bright reference images, is based on a teacher–student distillation mechanism with a momentum teacher, which allows the process of forming biometric representations to be stabilised. Additionally, aspects of keypoint stabilisation, loss gating based on detector confidence, and the balancing of a multi-component objective function are discussed. The proposed architecture provides a methodologically sound foundation for the development of a software suite aimed at achieving high accuracy in person recognition under complex lighting conditions.

Keywords: low light conditions, face detection, face recognition, identity preservation, common encoder, teacher-student distillation, loss gating.

Введение

Условия недостаточной освещённости ухудшают качество изображения не только визуально, но и на уровне признаков, которыми оперируют современные нейросетевые модели. Рост сенсорного шума, нестабильность экспозиции, локальные тени, снижение контраста и потеря высокочастотных деталей приводят к тому, что ключевые ориентиры лица определяются менее точно, а эмбединги, используемые для идентификации, становятся менее устойчивыми [1], [4]. Поэтому даже сильные модели детекции и распознавания, показавшие высокое качество на кадрах нормальной освещённости, в тёмных сценах демонстрируют выраженную деградацию точности [5], [7].

Традиционная каскадная схема, в которой улучшение изображения предшествует детекции и распознаванию, удобна для инженерной реализации, однако разрывает единую оптимизацию между этапами. Модуль улучшения в таком конвейере ориентируется прежде всего на визуальную выразительность изображения и не контролирует сохранение биометрически значимых деталей, тогда как последующие блоки вынуждены работать с входами из иного распределения, чем то, на котором они изначально обучались [5], [7]. В результате внешнее улучшение может не только не помочь, но и ухудшить распознавание, если сглаживаются тонкие текстуры, искажается локальная геометрия или изменяется микроконтраст в областях глаз, носа и рта.

В настоящей работе рассматривается сквозная постановка задачи, в которой улучшение изображения, детекция и распознавание объединяются в единый вычислительный граф. Базовая идея состоит в использовании общего энкодера признаков, к которому подключаются детекторные и распознавательные головы, а также механизмы сохранения идентичности, устойчивые к шуму, ошибкам раннего выравнивания и фотометрической деградации. Такой подход позволяет трактовать улучшение не как отдельную предобработку, а как часть общей процедуры формирования биометрически полезного представления [7].

Методы и принципы исследования

Исходное изображение рассматривается как входной тензор

$$x \in \mathbb{R}^{3 \times 3 \times H \times W} \quad (1)$$

где B обозначает размер пакета, а H и W задают пространственные размеры изображения. Общий E энкодер строит многоуровневое представление признаков, предназначенное для последующей совместной обработки задач улучшения, локализации и идентификации.

$$\{C_2, C_3, C_4, C_5\} \quad (2)$$

$$C_k \in \mathbb{R}^{B \times C \times \left(\frac{H}{2^k}\right) \times \left(\frac{W}{2^k}\right)} \quad (3)$$

Далее C_3, C_4, C_5 по уровням формируется пирамидальное представление признаков, обеспечивающее перенос информации между масштабами и позволяющее сохранить как локальные, так и более глобальные структуры лица.

$$\{P_3, P_4, P_5\} = \text{FPN}(C_3, C_4, C_5) \quad (4)$$

Полученное пирамидальное представление используется совместно в детекторе и распознавателе [6], [12]. Детектор D для каждого лица оценивает положение рамки, ключевые точки и меру уверенности, которая далее используется не только для локализации, но и для управления силой распознавательных потерь.

$$D(P_i) = \{b_i, l_i, s_i\}_{i=1}^N \quad (5)$$

Где b_i обозначает рамку;

$l = \{i\}$ — набор landmarks;

$s_i \in [0, 1]$ — confidence после подавления немаксимумов.

На следующем шаге выполняются выравнивание лица по ключевым точкам и извлечение фиксированного окна признаков, что позволяет привести входы распознавателя к согласованному геометрическому виду.

$$t_i = \text{ROIAlign}(P, A(b_{i,1})) \quad (6)$$

В этом выражении A обозначает аффинное выравнивание к каноническому виду, а t_i имеет фиксированный пространственный размер. После выравнивания распознавательная ветвь R формирует нормированный эмбединг, который затем используется в терминах идентичности и при необходимости в классификационной потере.

$$z_i = \text{norm}(R(t_i)) \quad (7)$$

$$z_i \in \mathbb{R}^d \quad (8)$$

2.1. Архитектурная постановка identity loss

Если эталонный эмбединг вычисляется непосредственно по тёмному входу, возникает логически замкнутый контур, в котором нестабильное представление используется как цель для собственной коррекции. В базовом варианте это можно записать следующим образом:

$$z_x = R(A(x, l_x)), \quad z_y = R(A(y, l_y)) \quad (9)$$

$$L_{iil} = 1 - \cos(z_x, z_y) \quad (10)$$

Здесь u обозначает улучшенное изображение. Ключевая проблема состоит в том, что при сильной деградации z_x сам по себе может быть шумным и смещённым, поэтому генератор получает ненадёжную целевую функцию и начинает подстраиваться под искажённый эталон [5].

Если доступны парные данные, то есть эталонное изображение нормальной освещённости x^{ref} , identityloss целесообразно определять по опорному изображению нормального качества:

$$z^{\text{ref}} = \text{norm}(R(A(x^{\text{ref}}, I^{\text{ref}}))), \quad z^y = \text{norm}(R(A(y, I^y))) \quad (11)$$

$$L_{\text{pair}}^i = 1 - \cos(z^{\text{ref}}, z^y) \quad (12)$$

Такой режим разрывает замкнутый круг, поскольку эталонный эмбединг извлекается из качественного изображения и тем самым задаёт более устойчивую цель для оптимизации [1].

При отсутствии парных данных предлагается использовать teacher student схему. В этой постановке teacher сеть R_τ обновляется по правилу экспоненциального скользящего среднего, что обеспечивает более плавную и устойчивую динамику формирования целевых представлений:

$$\theta_\tau \leftarrow \mu\theta_\tau + (1 - \mu)\theta_s, \quad \mu \in (0, 1) \quad (13)$$

Здесь θ_τ и θ_s обозначают параметры teacher и student соответственно. После обновления teacher формируются эмбединги, используемые для непарного варианта identity loss:

$$z_\tau = \text{norm}(R_\tau(A(y, I_y))), \quad z_s = \text{norm}(R_s(A(y, I_y))) \quad (14)$$

$$L_{\text{int}}^m = 1 - \cos(z_s, z_\tau) \quad (15)$$

Поскольку teacher ветвь по определению схемы не получает градиент, остановка градиента в данном случае выступает техническим следствием выбранной организации обучения, а не самостоятельным методом стабилизации [9].

Если доступны метки идентичности, в схему дополнительно включается ArcFace потеря:

$$L_{\text{arc}} = \text{ArcFaceLoss}(z, k) \quad (16)$$

Указанный терм усиливает разделимость классов в пространстве эмбедингов [5]. Если же метки идентичности отсутствуют, ArcFace потеря не используется, а сохранение идентичности обеспечивается исключительно через teacher student механизм и согласование представлений.

2.2. Стабилизация выравнивания и landmarks

На ранних этапах совместного обучения детектор и распознаватель ведут себя нестабильно, поскольку ошибки в landmarks немедленно отражаются на alignment и далее ухудшают эмбединги. Поэтому на этапе прогрева целесообразно использовать предобученный детектор D_0 как внешнего учителя для ключевых точек, а согласованность с его оценками вводить как дополнительный регуляризатор:

$$I^{(0)} = D_0 \quad (17)$$

Такое ограничение сдерживает отклонение текущих landmarks от стабильной опорной траектории и позволяет подключать identity потери только после того, как выравнивание становится достаточно надёжным для передачи корректного сигнала в распознаватель [6], [8].

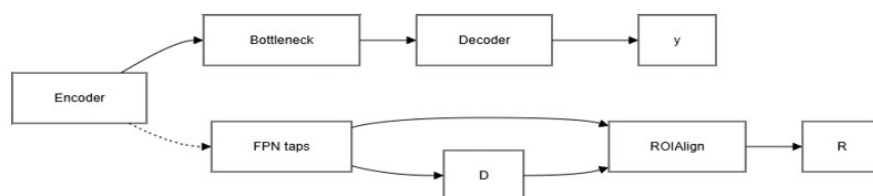


Рисунок 1 - Схема обучения

DOI: <https://doi.org/10.60797/IRJ.2026.168.4.1>

Сначала выполняется прогрев общего энкодера и ветви улучшения, затем стабилизируется детекция и ключевые точки, после чего включается совместное обучение с identity потерями.

Режимы данных и функция потерь

Под парным режимом понимается наличие пары «тёмное изображение, эталонное изображение» для одной сцены или одного лица. Под непарным режимом понимается отсутствие такого фотометрического эталона. Это различие не совпадает автоматически с наличием identity меток: метки личности могут присутствовать как в парном, так и в непарном сценарии. Такое разделение принципиально важно для корректного выбора супервизии.

Общая функция потерь модели записывается следующим выражением:

$$L = \lambda_{\text{enh}} L_{\text{enh}} + \lambda_{\text{rad}} L_{\text{rad}} + \lambda_{\text{det}} L_{\text{det}} + \lambda_{\text{arc}} L_{\text{arc}} + \lambda_{\text{id}} L_{\text{id}} + \lambda_{\text{cons}} L_{\text{cons}} \quad (18)$$

где, L_{cons} объединяет термины согласованности ключевых точек и признаков, ограничивая расхождение между ветвями и стабилизируя общее представление.

В парном режиме модуль улучшения естественно задаётся как сумма пиксельной, структурной и признаковой составляющих, что позволяет одновременно контролировать фотометрическое соответствие и сохранение признаков, значимых для распознавания:

$$L_{\text{enh}}^{\text{pair}} = \left\| y - x^{\text{ref}} \right\|_1 + \alpha \cdot (1 - \text{SSIM}(y, x^{\text{ref}})) + \beta \cdot L_{\text{feat}}(y, x^{\text{ref}}) \quad (19)$$

Коэффициенты α и β регулируют вклад структурной и признаковой частей. В непарном режиме вместо прямого эталона используются zero reference регуляризаторы, ориентированные на самосогласованность и корректное поведение улучшателя без явного фотометрического образца:

$$L_{\text{enh}} = L_{\text{enh}}^{\text{pair}}, \text{ если доступна пара } (x, x^{\text{ref}}) \quad (20)$$

$$L_{\text{enh}} = L_{\text{Z}}(y), \text{ если пары нет} \quad (21)$$

Аналогично adversarial часть включается только в том случае, если выбран GAN режим:

$$L_{\text{adv}} = 0, \text{ если GAN не используется} \quad (22)$$

Таким образом, непарность относится только к отсутствию фотометрического эталона и не отменяет возможность supervision для распознавания [2], [4]. Это позволяет разделять источник данных для улучшения и источник сигналов для идентификации без логических противоречий.

Балансировка многокомпонентной оптимизации

Поскольку отдельные функции потерь могут формировать разнонаправленные градиенты, статический подбор коэффициентов не всегда приводит к устойчивому обучению. Один из практических вариантов заключается в использовании обучаемых весов, связанных с гомоскедастической неопределённостью:

$$L = \sum_{i=1}^N \left(\frac{1}{2\sigma_i^2} L_i + \log(\sigma_i) \right) \quad (23)$$

В этом выражении σ_i обозначают обучаемые параметры, а индекс i соответствует отдельным задачам модели, например улучшению, детекции и распознаванию [10]. Такая форма позволяет автоматически перераспределять вклад задач по мере обучения. В качестве альтернативы может использоваться подход GradNorm, в котором контролируются нормы градиентов по общим параметрам:

$$G_i = \nabla_w (w_i L_i^2) \quad (24)$$

Здесь W обозначает набор общих параметров, а w_i — текущий вес задачи. Существенно, что сравниваются именно нормы градиентов, а не сами векторы производных, поскольку именно такая постановка позволяет корректно балансировать вклад задач в общее обновление параметров [10].

Результаты. Результатом проведённого анализа является непротиворечивая архитектурная схема, в которой улучшение, детекция и распознавание связаны через общий энкодер и согласованные функции потерь. По сравнению с каскадным конвейером предложенная постановка допускает прямое влияние задач детекции и распознавания на формирование общего пространства признаков. Это означает, что модуль улучшения оптимизируется не только по фотометрическим критериям, но и с учётом требований к локализации и устойчивости эмбеддингов [7].

Существенным результатом является устранение логической проблемы замкнутого круга identity loss. Показано, что использование эмбеддинга тёмного входа в качестве эталона методологически ненадёжно, тогда как парный эталон и teacher student схема формируют более устойчивые и интерпретируемые цели для оптимизации [1], [9]. Дополнительно обосновано, что ранняя стабилизация landmarks и staged обучение представляют собой не вспомогательные эвристики, а необходимый элемент архитектурной логики: без них распознаватель получает шумные и геометрически нестабильные входы [6], [8].

Отдельный результат связан с вычислительной организацией системы. Если головы детекции и распознавания подключены непосредственно к общему энкодеру, декодер улучшения может быть отключён на этапе инференса без потери основной функциональности, при условии, что обучение проводилось в том же режиме и градиенты от задач детекции и распознавания действительно влияли на общий энкодер. Это делает предложенную архитектуру применимой не только для офлайн обработки, но и для сценариев с ограниченным вычислительным бюджетом.

Дополнительным результатом является повышение отказоустойчивости за счёт гейтинга по уверенности детектора. Пусть s_i обозначает confidence для i -й рамки после подавления немаксимумов. Тогда вес распознавательной части может задаваться следующей функцией:

$$g(s_i) = \sigma(k(s_i - \tau)), \quad k > 0 \quad (25)$$

Соответствующий гейтированный identity loss принимает вид:

$$L_{\text{id}}^{\text{gate}} = \sum_i g(s_i)(1 - \cos(z_{s,i}, z_{\tau,i})) \quad (26)$$

Аналогичным образом определяется и гейтированная ArcFace потеря:



$$L_{arc}^{gate} = \sum_i g(s_i) L_{arc,i} \quad (27)$$

Такая форма ослабляет влияние распознавания на ранних и слабодетектируемых примерах, не отключая распознаватель полностью и не разрушая совместную оптимизацию.

Предложенная схема показывает, что ключевые инженерные риски в задачах low light распознавания лиц определяются не столько выбором конкретного backbone, сколько согласованием источников супервизии и моментом включения распознавательных термов. Если identity loss привязан к нестабильному эмбедингу тёмного входа, система получает ошибочную цель вне зависимости от мощности генератора. Напротив, при использовании согласованных признаков общего энкодера и при подключении identity термов только после стабилизации landmarks архитектура становится существенно более устойчивой как с методологической, так и с вычислительной точки зрения.

Не менее важна и вычислительная организация модели. Shared encoder с прямыми подключениями к головам детекции и распознавания уменьшает избыточную стоимость обработки, а возможность отключения декодера на инференсе делает схему пригодной для практических систем, где критичны время отклика и объём памяти. Дополнительный гейтинг распознавательных потерь по уверенности детектора естественным образом учитывает качество наблюдения и снижает риск разрушительных градиентов на ранних стадиях обучения.

Дополнительное методологическое значение предложенной архитектуры состоит в том, что она позволяет явно развести фотометрическую и биометрическую составляющие задачи, не разрывая их на уровне обучения. В классических каскадах между этими составляющими обычно отсутствует обратная связь: модуль улучшения стремится восстановить визуальную разборчивость кадра, а распознаватель лишь пассивно принимает полученный результат. В рассматриваемой схеме, напротив, улучшение подчиняется требованиям последующих задач и тем самым оптимизируется в контексте конечной цели. Это особенно важно для сцен с неравномерным освещением, локальными пересветами, тенями и сенсорным шумом, где визуально «приятное» восстановление не всегда совпадает с сохранением идентичности значимых микропризнаков. Следовательно, даже без проведения экспериментальной части уже на уровне архитектурного анализа можно утверждать, что согласование источников супервизии и последовательности включения потерь является критическим условием устойчивой работы системы.

Практическая перспектива дальнейшей верификации такой модели связана с поэтапной экспериментальной проверкой на наборах данных, содержащих как контролируемые, так и естественные низкоосвещённые сцены. На первом этапе необходимо подтвердить устойчивость детекции и корректность landmarks после прогрева и регуляризации, на втором этапе — оценить влияние парного и непарного режимов на сохранение идентичности, а на третьем этапе — сопоставить качество распознавания при включённом и отключённом декодере улучшения на инференсе. Подобная стратегия не меняет предложенную архитектурную логику, но позволяет последовательно проверить те допущения, которые были сформулированы в данной статье на теоретическом уровне. Тем самым предложенная схема выступает не только как концептуальная модель, но и как практически пригодный каркас для построения последующих экспериментальных протоколов.

Заключение

Предложенная сквозная архитектура объединяет улучшение изображения, детекцию и распознавание лиц в единый вычислительный контур и устраняет ключевые противоречия, характерные для каскадных схем. Основной вклад работы состоит в том, что улучшение рассматривается не как внешняя предобработка, а как часть процесса формирования биометрически полезного представления. Благодаря этому фотометрическая коррекция связывается с требованиями к точности детекции и к устойчивости эмбедингов.

Показано, что наиболее критичным риском является использование нестабильного тёмного эмбединга как эталона для identity loss. Для устранения этого риска предложены два согласованных режима: первый основан на парном эталонном изображении, второй использует teacher student дистилляцию с momentum teacher. Кроме того, обоснована необходимость ранней стабилизации landmarks, адаптивного взвешивания многокомпонентной функции потерь и гейтинга распознавательных термов по уверенности детектора.

С практической точки зрения наибольший интерес представляет вариант с общим энкодером и прямыми подключениями голов детекции и распознавания, поскольку он позволяет отключать декодер улучшения на этапе инференса и тем самым снижать вычислительную нагрузку. Следовательно, сформулированная архитектурная схема может служить основой для дальнейшей программной реализации, экспериментальной верификации и прикладного внедрения в задачах распознавания лиц при низкой освещённости.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Chen C. Learning to see in the dark / C. Chen, Q. Chen Q., J. Xu et al. // In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2018. — № 16. — DOI: 10.1109/CVPR.2018.00347



2. Guo C. Zero-reference deep curve estimation for low-light image enhancement / C. Guo, C. Li, J. Guo // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). — 2020. — DOI: 10.1109/CVPR42600.2020.00185.
3. Ma L. Toward fast, flexible, and robust low-light image enhancement / L. Ma, T. Ma, R. Liu // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2022. — DOI: 10.1109/CVPR52688.2022.00555.
4. Jiang Y. EnlightenGAN: deep light enhancement without paired supervision / Y. Jiang, X. Gong, D. Liu // IEEE Transactions on Image Processing. — 2021. — № 30. — P. 2340-23-49. — DOI: 10.1109/TIP.2021.3051462
5. Deng J. ArcFace: additive angular margin loss for deep face recognition / J. Deng, J. Guo, N. Xue et al. // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2019. — № 22. — DOI: 10.1109/CVPR.2019.00482
6. Deng J. RetinaFace: single-shot multi-level face localisation in the wild / J. Deng, J. Guo, E. Ververas et al. // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2020. — № 6. — DOI: 10.1109/CVPR42600.2020.00525
7. Wang W. HLA-Face: joint high-low adaptation for low-light face detection / W. Wang, W. Yang, J. Liu // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2021. — № 77. — DOI: 10.1109/CVPR46437.2021.01593
8. Визильтер Ю.В. Одноэтапный детектор лиц и особых точек на цифровых изображениях / Ю.В. Визильтер, В.С. Горбацевич, А.С. Моисеенко // Компьютерная оптика. — 2020. — № 44 (4). — С. 589–595. — DOI: 10.18287/2412-6179-CO-674
9. Tarvainen A. Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results / A. Tarvainen, H. Valpola // arXiv. — 2017. — DOI: 10.48550/arXiv.1703.01780
10. Kendall A. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics / A. Kendall, Y. Gal, R. Cipolla // Proceedings of the IEEE/CVF. — 2018. — № 21. — DOI: 10.1109/CVPR.2018.00781

Список литературы на английском языке / References in English

1. Chen C. Learning to see in the dark / C. Chen, Q. Chen Q., J. Xu et al. // In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2018. — № 16. — DOI: 10.1109/CVPR.2018.00347
2. Guo C. Zero-reference deep curve estimation for low-light image enhancement / C. Guo, C. Li, J. Guo // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). — 2020. — DOI: 10.1109/CVPR42600.2020.00185.
3. Ma L. Toward fast, flexible, and robust low-light image enhancement / L. Ma, T. Ma, R. Liu // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2022. — DOI: 10.1109/CVPR52688.2022.00555.
4. Jiang Y. EnlightenGAN: deep light enhancement without paired supervision / Y. Jiang, X. Gong, D. Liu // IEEE Transactions on Image Processing. — 2021. — № 30. — P. 2340-23-49. — DOI: 10.1109/TIP.2021.3051462
5. Deng J. ArcFace: additive angular margin loss for deep face recognition / J. Deng, J. Guo, N. Xue et al. // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2019. — № 22. — DOI: 10.1109/CVPR.2019.00482
6. Deng J. RetinaFace: single-shot multi-level face localisation in the wild / J. Deng, J. Guo, E. Ververas et al. // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2020. — № 6. — DOI: 10.1109/CVPR42600.2020.00525
7. Wang W. HLA-Face: joint high-low adaptation for low-light face detection / W. Wang, W. Yang, J. Liu // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2021. — № 77. — DOI: 10.1109/CVPR46437.2021.01593
8. Vizil'ter Yu.V. Odnoe'tapny'j detektor licz i osoby'x toчек na cifrovyy'x izobrazheniyax [One-step detector of faces and special points on digital images] / Yu.V. Vizil'ter, V.S. Gorbacevich, A.S. Moiseenko // Computer optics. — 2020. — № 44 (4). — P. 589–595. — DOI: 10.18287/2412-6179-CO-674 [in Russian]
9. Tarvainen A. Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results / A. Tarvainen, H. Valpola // arXiv. — 2017. — DOI: 10.48550/arXiv.1703.01780
10. Kendall A. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics / A. Kendall, Y. Gal, R. Cipolla // Proceedings of the IEEE/CVF. — 2018. — № 21. — DOI: 10.1109/CVPR.2018.00781