# ГИБРИДНЫЕ СИСТЕМЫ РАСПОЗНАВАНИЯ ЧЕЛОВЕКА ПО ИЗОБРАЖЕНИЮ ГЛАЗА

Хамидов Эльнур Хамидович
*Национальный исследовательский университет «Ташкентский институт инженеров ирригации и механизации сельского хозяйства»*
elnur.khamidov1587@gmail.com

***Annotatsiya.*** *Maqolada inson ko'z tasviri bo'yicha tanib olishning gibrid tizimlari tahlil qilingan. 2022–2026-yillarda chop etilgan ilmiy ishlar asosida gibridlashning uch turi ajratilgan: arxitekturaviy (konvolyutsion neyron tarmoqlar va transformerlarni birlashtirish), belgilar darajasidagi (an'anaviy deskriptorlar va chuqur o'rganish belgilarini birlashtirish) hamda ko'z sohalari bo'yicha (ko'z rangdor pardasi, ko'z atrofi sohasi va skleradan birgalikda foydalanish). Mavjud yondashuvlarning afzalliklari va cheklovlari, jumladan, birlashtirish sxemalarining statikligi va natijalarning sust taqqoslanuvchanligi ko'rsatib berilgan. Tasvir sifatining fazoviy xaritasi o'zaro diqqat mexanizmini va baholar darajasida moslashuvchan birlashtirishni boshqaradigan gibrid tizimning umumlashtirilgan arxitekturasi taklif etilgan.*

***Kalit so'zlar:*** *biometrik identifikatsiya, ko'z rangdor pardasi, ko'z atrofi sohasi, gibrid tizimlar, konvolyutsion neyron tarmoqlar, transformerlar, ma'lumotlarni birlashtirish, tasvir sifati.*

***Аннотация.*** *В работе проанализированы гибридные системы распознавания человека по изображению глаза. На основе публикаций 2022–2026 гг. выделены три типа гибридизации: архитектурная (сочетание свёрточных нейронных сетей и трансформеров), гибридизация признаков (объединение классических дескрипторов и глубоких признаков) и многообластная (совместное использование радужной оболочки, периокулярной области и склеры). Показаны преимущества и ограничения существующих подходов, в том числе статичность схем слияния и слабая сопоставимость результатов. Предложена обобщённая архитектура гибридной системы, в которой пространственная карта качества изображения управляет механизмом перекрёстного внимания и адаптивным слиянием на уровне оценок.*

***Ключевые слова:*** *биометрическая идентификация, радужная оболочка глаза, периокулярная область, гибридные системы, свёрточные нейронные сети, трансформеры, слияние данных, качество изображения.*

***Abstract.*** *The paper analyzes hybrid systems for human recognition from eye images. Based on studies published in 2022–2026, three types of hybridization are identified: architectural (combining convolutional neural networks and transformers), feature-level (fusing handcrafted descriptors with deep features), and multi-region (jointly using the iris, the periocular region and the sclera). The advantages and limitations of existing approaches are discussed, including static fusion schemes and the weak comparability of reported results. A generalized architecture of a hybrid system is proposed, in which a spatial image-quality map controls cross-attention and adaptive score-level fusion.*

***Keywords:*** *biometric identification, iris, periocular region, hybrid systems, convolutional neural networks, transformers, information fusion, image quality.*

**Введение.** Распознавание человека по изображению глаза относится к наиболее точным методам биометрической идентификации: текстура радужной оболочки уникальна и стабильна в течение жизни, а окружающая её периокулярная область и сосудистый рисунок склеры несут дополнительную различительную информацию. Однако при переходе от контролируемой съёмки в ближнем инфракрасном (NIR) диапазоне к неконтролируемым условиям — съёмке в видимом спектре (VIS), на расстоянии, мобильными устройствами и шлемами виртуальной реальности — качество изображений резко снижается из-за размытия, окклюзий веками и ресницами, бликов и отклонения взгляда [1, 2]. В таких условиях системы, построенные на одном типе модели или на одной области глаза, теряют устойчивость.

Ответом на эту проблему стали гибридные системы распознавания, в которых объединяются разнородные компоненты: модели разных классов, признаки разной природы или данные разных областей глаза. Систематические обзоры показывают, что такие системы, как правило, превосходят одномодальные решения, причём слияние может выполняться на уровне данных, признаков, оценок или решений, а наиболее распространённым остаётся слияние на уровне оценок [2, 3]. Цель данной работы — систематизировать гибридные подходы к распознаванию по изображению глаза на основе публикаций 2022–2026 гг., выявить их общие ограничения и предложить обобщённую архитектуру гибридной системы с управлением по качеству изображения.

**Классификация гибридных систем.** По результатам анализа литературы гибридные системы предлагается разделить на три группы (табл. 1).

**Таблица 1 – Примеры гибридных систем распознавания по изображению глаза (2023–2026 гг.)**

| Работа | Тип гибридизации | Компоненты системы | Данные и результат |
|---|---|---|---|
| [4] | Архитектурная | CNN + ViT | три открытые базы; точность до 93,66 % |
| [7] | Архитектурная (конвейерная) | детекция глаза и радужки, сегментация, Swin Transformer | CASIA-Iris-Thousand, -Lamp, -Interval; точность 95,14–99,56 % |
| [8] | Архитектурная (сегментация) | Swin Transformer + CNN, перекрёстное внимание | UBIRIS.v2, IITD, CASIA-Iris-Thousand, MMU; Dice 0,949–0,952 |
| [9] | Признаковая | MEEMD + LBP + VGG16 / ResNet-152 | IITD 98 %, CASIA 97 %, UBIRIS.v2 97,3 % |
| [10] | Признаковая | вейвлет-фильтр Чебышёва + CNN | CASIA-IrisV4, UBIRIS.v2, MMU; точность до 98,97 % |
| [11] | Признаковая | фильтры Габора + карты вероятности окклюзий | CASIA-Iris-Thousand, IITD; EER 0,51 % |
| [13] | Признаковая (обнаружение атак) | адаптивные бинарные шаблоны + ViT | Notre Dame, IIITD-WVU, NDCLD15; ACER 1,52–1,97 % |
| [15] | Многообластная | радужка + периокулярная область, CNN | CASIA-Iris-Thousand; точность 93,64 → 99,80 % |
| [16] | Многообластная | радужка + склера, DWT, сиамская сеть | точность 98,5 % (бимодальный режим) |
| [17] | Многообластная, с учётом качества | DINOv2 + LoRA, слияние шаблонов по качеству | UBIPr, UTIRIS; EER 1,2 % |

*Примечание:* показатели получены в разных протоколах оценки и не являются непосредственно сопоставимыми.

*Архитектурная гибридизация* сочетает свёрточную нейронную сеть (CNN), эффективно извлекающую локальную текстуру, с трансформером, моделирующим глобальные зависимости между фрагментами изображения. Комбинированная модель CNN + ViT обеспечила точность распознавания до 93,66 % на трёх открытых базах [4], а её встраивание в сиамскую архитектуру позволило верифицировать новых субъектов по одному эталонному образцу [5]. Работа IrisFormer показала, что трансформер раскрывает свой потенциал для радужной оболочки лишь при учёте её специфики: относительного позиционного кодирования, горизонтальных сдвигов при обучении для инвариантности к повороту и случайного маскирования токенов для

устойчивости к окклюзиям [6]. Конвейер SwinIris объединяет детекцию глаза, детекцию и сегментацию радужки и классификацию на основе Swin Transformer [7]. Принцип гибридизации распространяется и на сегментацию: в SwinIrisNet ветви Swin Transformer и CNN объединяются модулем перекрёстного внимания, что обеспечивает коэффициент Dice 0,949–0,952 на базах UBIRIS.v2, IITD, CASIA-Iris-Thousand и MMU [8].

*Гибридизация признаков* объединяет классические (handcrafted) дескрипторы с глубокими признаками. Классические дескрипторы вносят в модель априорные знания о текстуре и сохраняют работоспособность при малых обучающих выборках, тогда как глубокие признаки обладают высокой обобщающей способностью. Примерами служат объединение статистических признаков MEEMD, локальных бинарных шаблонов (LBP) и признаков сетей VGG16 и ResNet-152 [9]; сочетание двумерного вейвлет-фильтра Чебышёва с CNN [10]; банк многомасштабных фильтров Габора с адаптивным взвешиванием признаков по картам вероятности окклюзий [11]; встраивание фильтров Эрмита в нейросетевой детектор ключевых точек, не требующий масок сегментации [12]. Тот же принцип эффективен и для обнаружения атак предъявления: адаптивные бинарные карты признаков в сочетании с ViT снижают средний коэффициент ошибок классификации (ACER) до 1,52–1,97 % [13].

*Многообластная гибридизация* использует совместно несколько областей одного изображения глаза: радужную оболочку, периокулярную область, склеру, а также динамику взгляда. Периокулярная область менее требовательна к разрешению и кооперативности пользователя и поэтому компенсирует снижение качества радужки. В системах распознавания на расстоянии оценка качества радужки используется непосредственно как вес при слиянии с периокулярными признаками [14]; добавление периокулярной информации к CNN-распознаванию радужки повысило точность на базе CASIA-Iris-Thousand с 93,64 до 99,80 % [15]. Слияние признаков радужки и склеры на основе дискретного вейвлет-преобразования дало точность 98,5 %, превысив одномодальные варианты [16]. Для межспектрального периокулярного распознавания применяются базовые модели (foundation models): DINOv2 с адаптацией LoRA и слиянием шаблонов с учётом качества достигает EER 1,2 % [17]. Слияние периокулярных изображений с движениями глаз на выборке из 9202 субъектов стабильно превосходит каждую модальность в отдельности [18].

**Ограничения существующих подходов.** Анализ выявил ряд общих проблем. Во-первых, результаты слабо сопоставимы: многие работы приводят

точность замкнутой классификации на небольших выборках вместо показателей верификации (EER, FNMR при фиксированном FMR) в открытом протоколе, а описание разбиения данных часто недостаточно для воспроизведения эксперимента [19]. Во-вторых, в большинстве систем веса слияния фиксированы и не учитывают качество конкретного изображения, хотя именно качество определяет, какой компонент надёжнее в данный момент [14, 17]. В-третьих, гибридные модели с трансформерами вычислительно затратны, что затрудняет их применение на мобильных и встраиваемых устройствах; это стимулирует разработку облегчённых архитектур, например сочетания разделимых свёрток с классификатором SVM [20]. Наконец, многие системы сохраняют жёсткую зависимость от точности сегментации и нормализации, ошибки которых распространяются на все последующие этапы [2].

**Обобщённая архитектура гибридной системы с управлением по качеству.** С учётом выявленных ограничений предлагается обобщённая архитектура (рис. 1), объединяющая все три вида гибридизации и использующая пространственную карту качества изображения в качестве управляющего сигнала.

На первом этапе гибридная сеть сегментации (CNN + Transformer) выделяет зрачок, радужную оболочку, склеру и периокулярную область. Параллельно строится карта качества:

$$q(x, y) = [1 - o(x, y)] \cdot \varphi(v(x, y)) \cdot \psi(c(x, y)), \quad (1)$$

где $o(x, y)$ — маска окклюзий и бликов; $v(x, y)$ — локальная дисперсия лапласиана (мера резкости); $c(x, y)$ — локальный контраст; $\varphi$, $\psi$ — функции нормировки к интервалу [0, 1]. На начальном этапе карта вычисляется классическими методами и не требует обучения.

В ветви радужной оболочки CNN с круговыми фильтрами Габора извлекает локальную текстуру, а трансформер формирует глобальный контекст. Признаки объединяются перекрёстным вниманием, модулированным качеством:

$$\mathrm{Attn} = \mathrm{softmax}(Q_C K_T^{\top} / \sqrt{d} + \lambda \cdot \log(q_p + \varepsilon)) \cdot V_T, \quad (2)$$

где $Q_C$ — запросы, сформированные из токенов CNN; $K_T$, $V_T$ — ключи и значения токенов трансформера; $d$ — размерность признаков; $q_p$ — среднее значение карты качества в пределах фрагмента $p$; $\lambda$ — обучаемый коэффициент; $\varepsilon$ — малая константа. Фрагменты низкого качества получают большой отрицательный сдвиг и практически исключаются из внимания, тогда как в стандартном механизме они участвуют наравне с остальными.

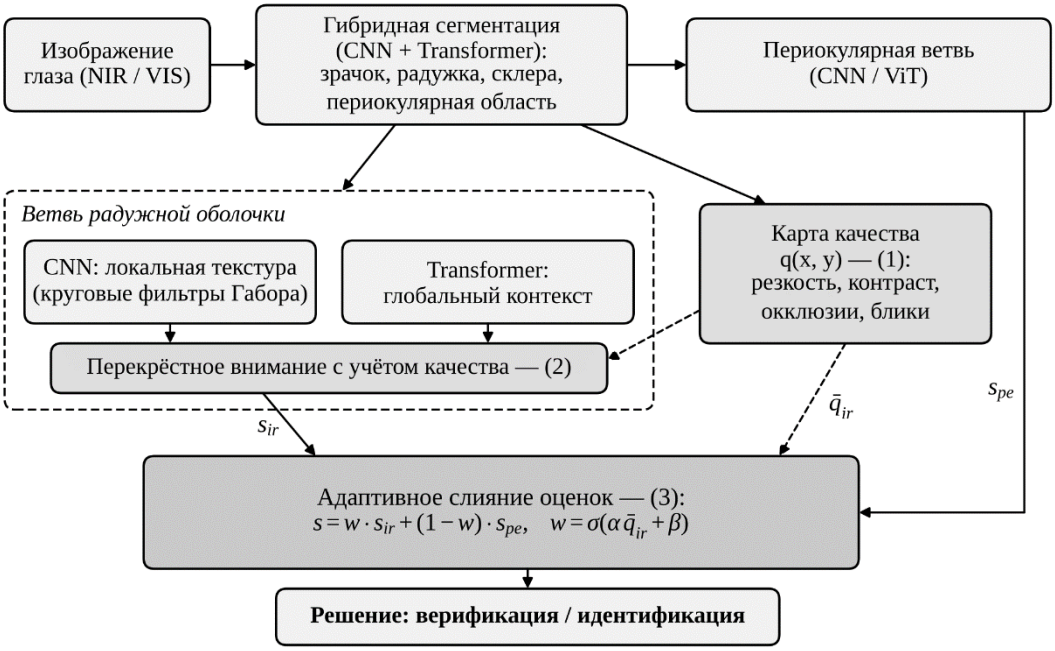


***Рисунок 1 – Обобщённая архитектура гибридной системы распознавания по изображению глаза с управлением по качеству***

Для компенсации поворота глаза при сравнении шаблонов используются циклические сдвиги вдоль угловой координаты нормализованного изображения. Итоговое решение принимается адаптивным слиянием оценок сходства радужной и периокулярной ветвей:

$$s = w \cdot s_{ir} + (1 - w) \cdot s_{pe}, \quad w = \sigma(\alpha \cdot \bar{q}_{ir} + \beta), \tag{3}$$

где $s_{ir}$, $s_{pe}$ — оценки сходства радужной и периокулярной ветвей; $\bar{q}_{ir}$ — средний показатель качества области радужки; $\sigma$ — сигмоидная функция; $\alpha$, $\beta$ — параметры, настраиваемые на валидационной выборке. При высоком качестве радужки решение определяется преимущественно ею, а при деградации вес автоматически смещается в сторону периокулярной области.

Предложенная архитектура обладает следующими свойствами: качество учитывается на уровне отдельных фрагментов, а не изображения в целом; при деградации одной из областей точность снижается плавно, без отказа системы; явная карта качества делает решения интерпретируемыми. Экспериментальную проверку предполагается провести на базах CASIA-Iris-Thousand, IITD и UBIRIS.v2, упорядоченных по степени деградации изображений, с оценкой показателей EER и FNMR при FMR = $10^{-3}$ во внутри- и межбазовых протоколах.

**Заключение.** Гибридные системы стали основным направлением развития распознавания человека по изображению глаза. Выделены три вида гибридизации — архитектурная, на уровне признаков и многообластная, — каждый из которых по-своему повышает устойчивость к неконтролируемым

условиям съёмки. Главными ограничениями остаются статичность схем слияния, вычислительная сложность и слабая сопоставимость результатов. Предложенная архитектура с управлением по пространственной карте качества направлена на устранение первого из этих ограничений; дальнейшие исследования будут посвящены её экспериментальной проверке и разработке облегчённой версии для мобильных устройств.

## FOYDALANILGAN ADABIYOTLAR

1. Nguyen K., Proença H., Alonso-Fernandez F. Deep Learning for Iris Recognition: A Survey // ACM Computing Surveys. – 2024. – Vol. 56, No. 9. – DOI: 10.1145/3651306.

2. Yu Y. [et al.] Segmentation and Multimodal Fusion in Ocular Biometric Recognition: A Review // Medinformatics. – 2026. – DOI: 10.47852/bonviewmedin62029908.

3. Es-Sobbahi H. [et al.] Multimodal Biometrics: A Review of Handcrafted and AI-Based Fusion Approaches // IET Biometrics. – 2025. – Art. 5055434. – DOI: 10.1049/bme2/5055434.

4. Latif S.A. [et al.] An Efficient Iris Recognition Technique using CNN and Vision Transformer // Journal of Advanced Research in Applied Sciences and Engineering Technology. – 2023. – Vol. 34, No. 2. – P. 235–245. – DOI: 10.37934/araset.34.2.235245.

5. Latif S.A. [et al.] Hybrid CNN-ViT integration into Siamese networks for robust iris biometric verification // Bulletin of Electrical Engineering and Informatics. – 2025. – Vol. 14, No. 4. – DOI: 10.11591/eei.v14i4.9182.

6. Sun X. [et al.] IrisFormer: A Dedicated Transformer Framework for Iris Recognition // IEEE Signal Processing Letters. – 2025. – Vol. 32. – DOI: 10.1109/LSP.2024.3522856.

7. Gao R. [et al.] On Designing a SwinIris Transformer Based Iris Recognition System // IEEE Access. – 2024. – Vol. 12. – DOI: 10.1109/ACCESS.2024.3369035.

8. Oteko T.L. [et al.] SwinIrisNet: A Hybrid Deep Learning Framework for Robust Iris Segmentation // Applied Sciences. – 2026. – Vol. 16, No. 17. – Art. 8892. – DOI: 10.3390/app16178892.

9. Louriga Z.E. [et al.] A Hybrid Multi-Domain Feature Fusion Model Integrating MEEMD and Dual CNN for Iris Recognition // Machine Learning and Knowledge Extraction. – 2026. – Vol. 8, No. 4. – Art. 111. – DOI: 10.3390/make8040111.

10. Aliyu M.G. [et al.] An Enhanced Iris Recognition using Deep Learning and 2D Chebyshev Wavelet Filter // Journal of Information and Communication Technology. – 2026. – Vol. 25, No. 3. – P. 98–123. – DOI: 10.32890/jict2026.25.3.5.

11. Tanna R. [et al.] OcclusionNetPlusPlus: A Multi-Scale Similarity Network with Adaptive Occlusion Detection for Robust Iris Recognition // International Journal of Cognitive Computing in Engineering. – 2025. – DOI: 10.1016/j.ijcce.2025.09.002.

12. Andreeva E. [et al.] Iris image key points extraction based on handcrafted features in neural network // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2025. – Vol. XLVIII-2/W9-2025. – DOI: 10.5194/isprs-archives-XLVIII-2-W9-2025-7-2025.

13. Sharma D. [et al.] Cascading adaptive binary image feature maps with vision transformer for iris spoof detection // Applied Soft Computing. – 2025. – Art. 112713. – DOI: 10.1016/j.asoc.2025.112713.

14. Yan Z. [et al.] Boosting Multi-modal Ocular Recognition via Spatial Feature Reconstruction and Unsupervised Image Quality Estimation // Machine Intelligence Research. – 2024. – Vol. 21. – DOI: 10.1007/s11633-023-1415-y.

15. Anegsiripong K. [et al.] Personal Identification Based on Fusion of Iris and Periocular Information Using Convolutional Neural Networks // 62nd Annual Conference of the Society of Instrument and Control Engineers (SICE). – 2023. – DOI: 10.23919/SICE59929.2023.10354096.

16. Adeniyi J.K. [et al.] Fusion of Siamese Network-Based Sclera and Iris Detection: A Multimodal Biometrics Approach Using a Sclera Detection Tracing Algorithm // IET Information Security. – 2026. – Art. 5569382. – DOI: 10.1049/ise2/5569382.

17. Prasad D. [et al.] Robust Cross-Spectral Periocular Biometric Recognition using LoRA-Adapted DINOv2 and Quality-Aware Fusion // IETE Journal of Research. – 2026. – P. 1–11. – DOI: 10.1080/03772063.2026.2695248.

18. Lohr D. [et al.] Ocular Authentication: Fusion of Gaze and Periocular Modalities // 2025 IEEE International Joint Conference on Biometrics (IJCB). – 2025. – DOI: 10.1109/IJCB65343.2025.11411028.

19. Alrikabi J.M. A Combination Approach for Iris Recognition Based on Convolutional Neural Networks // Iraqi Journal for Computers and Informatics. – 2026. – Vol. 52, No. 1. – P. 368–379. – DOI: 10.25195/ijci.v52i1.778.

20. Patil C.M. [et al.] Lightweight deep convolutional neural network architecture for real-time iris recognition in resource-constrained environments // Engineering Research Express. – 2026. – Vol. 8, No. 1. – Art. 015210. – DOI: 10.1088/2631-8695/ae30cb.