

O'ZGARUVCHAN YORUG'LIK VA FON SHAROITLARIDA YUZNI ANIQLASH VA TANIB OLIISHNING GIBRID MODEL I

Saidqodirov Xumoyunxon Yashnarjon o'g'li
Xusanboyeva Farzonaxon Ismoiljon qizi

Toshkent axborot texnologiyalari universiteti, Toshkent, O'zbekiston DIF 310-23 guruh talabasi,
DIF 319-24 guruh talabasi,
xumoyun.saidqodirov@gmail.com,
xusanboyevafarzonaxon@gmail.com

Annotatsiya: Ushbu maqolada o'zgaruvchan yorug'lik va fon sharoitlarida yuzni aniqlash va tanib olish uchun mo'ljallangan gibrid model taqdim etiladi. Dolzarbligi shundaki, real dunyo ilovalarida mavjud tizimlar past yoritilganlik, murakkab fonlar va okkluziya sharoitlarida aniqligini yo'qotadi. Muammoni hal qilish uchun Retinex asosidagi yorug'lik normalizatsiyasi, HOG, LBP va CNN xususiyatlarini birlashtiradigan ko'p kanalli yondashuv hamda og'irlikli ansambl klassifikatsiyasi taklif etiladi. Natijada LFW ma'lumotlar to'plamida 94.7% aniqlik va 31 FPS tezligiga erishildi. Tadqiqotning ilmiy yangiligi — yorug'lik normalizatsiyasi bilan ko'p kanalli gibrid arxitekturani birlashtirishdan iborat bo'lib, bu mavjud usullarga nisbatan 12-15% aniqlikni oshiradi. Amaliy ahamiyati — xavfsizlik tizimlari, bank autentifikatsiyasi va tibbiy diagnostika sohalarida qo'llanilishi mumkin.

Kalit so'zlar: yuzni aniqlash, yuzni tanib olish, gibrid model, chuqur o'rganish, CNN, HOG, LBP, yorug'lik normalizatsiyasi, real vaqt qayta ishlash.

Аннотация: В данной статье представлена гибридная модель для обнаружения и распознавания лиц в условиях изменяющегося освещения и фона. Актуальность обусловлена тем, что существующие системы теряют точность при слабом освещении, сложных фонах и окклюзии. Предложен многоканальный подход, объединяющий нормализацию освещения на основе Retinex, дескрипторы HOG, LBP и CNN, а также взвешенную ансамблевую классификацию. Достигнута точность 94.7% и скорость 31 FPS на наборе данных LFW. Научная новизна — совместное применение нормализации освещения и многоканальной гибридной архитектуры, повышающее точность на 12-15% по сравнению с существующими методами.

Ключевые слова: обнаружение лиц, распознавание лиц, гибридная модель, глубокое обучение, CNN, HOG, LBP, нормализация освещения, обработка в реальном времени.



Abstract: *This paper presents a hybrid model for face detection and recognition under variable lighting and background conditions. The relevance is determined by the fact that existing systems lose accuracy under low illumination, complex backgrounds, and occlusion. A multi-channel approach combining Retinex-based illumination normalization, HOG, LBP and CNN features, along with weighted ensemble classification, is proposed. As a result, 94.7% accuracy and 31 FPS processing speed were achieved on the LFW dataset. The scientific novelty lies in the joint application of illumination normalization and multi-channel hybrid architecture, which improves accuracy by 12-15% compared to existing methods.*

Keywords: *face detection, face recognition, hybrid model, deep learning, CNN, HOG, LBP, illumination normalization, real-time processing.*

KIRISH

Yuzni aniqlash va tanib olish (Face Detection and Recognition — FDR) — sun'iy intellekt va kompyuter ko'rishining eng faol rivojlanayotgan yo'nalishlaridan biri hisoblanadi. Xavfsizlik tizimlari, bank autentifikatsiyasi, aqlli qurilmalar va tibbiy diagnostika kabi sohalarda bu texnologiyaning ahamiyati tobora ortib bormoqda [1, 2].

Amaliy muhitlarda tizimlar ko'pincha quyidagi muammolarga duch keladi: o'zgaruvchan yorug'lik (past yoritilganlik, ko'r-ko'rona nur), murakkab va dinamik fonlar, yuz pozasidagi o'zgarishlar (90 darajagacha og'ish) hamda okkluziya — yuzning qisman berkitilishi. Bu omillar mavjud tizimlarning aniqligi va ishonchliligini sezilarli pasaytiradi [3].

Mavjud yuzni tanib olish tizimlari asosan ikki yondashuvga asoslanadi: an'anaviy belgi ajratish usullari (HOG, LBP, Eigenfaces) va neyron tarmoqqa asoslangan chuqur o'rganish (CNN, VGG, ResNet). An'anaviy usullar tezkor va kam resurs talab qiladi, ammo o'zgaruvchan sharoitlarda aniqligi pasayadi. Chuqur o'rganish usullari yuqori aniqlik beradi, lekin katta hajmli o'quv ma'lumotlari va hisoblash quvvatini talab qiladi [4].

Tadqiqotning maqsadi — ikkala yondashuvning kuchli tomonlarini birlashtirgan gibrid model yaratish orqali real dunyo sharoitlarida yuqori aniqlik, tezlik va ishonchlilikni bir vaqtda ta'minlash. Ilmiy yangilik sifatida Retinex asosidagi yorug'lik normalizatsiyasi bilan ko'p kanalli xususiyat ajratish va og'irlikli ansambl klassifikatsiyasining birlashtirilishi taklif etiladi.

ADABIYOTLAR SHARHI

Yuzni aniqlash sohasidagi ilk muhim ishlardan biri Viola va Jones [5] tomonidan 2001-yilda taklif etilgan Haar kaskad detektor hisoblanadi. Real vaqt rejimida ishlash qobiliyatiga ega bo'lgan ushbu usul ko'plab tizimlarga asos bo'ldi. Dalal va Triggs [6] 2005-yilda HOG (Histogram of Oriented Gradients) xususiyatlarini taqdim etdi.



Ojala va Pietikäinen [7] tomonidan ishlab chiqilgan LBP (Local Binary Patterns) tavsifi to'qima tahlili va yuzni tanib olishda keng qo'llaniladi. LBP yorug'lik o'zgarishlariga nisbatan ma'lum darajada chidamli bo'lib, hisobiy resurslarni kam talab qilishi bilan ajralib turadi.

Chuqur o'rganish sohasida DeepFace [8] va FaceNet [9] tizimlari inson darajasiga yaqin aniqlikka erishdi. Biroq, bu modellar yorug'lik va fon o'zgarishlariga sezgir bo'lib, amaliy sharoitlarda aniqlik pasayishi kuzatiladi [10]. Gibril yondashuvlar [11, 12] an'anaviy usullar va chuqur o'rganishni birlashtirishning samaradorligini ko'rsatmoqda, ammo yorug'lik normalizatsiyasi masalasi hali to'liq hal etilmagan.

METODOLOGIYA

3.1. Tizim arxitekturasini

Taklif etilayotgan gibril model uchta asosiy moduldan iborat: (1) Oldindan qayta ishlash moduli — Retinex va CLAHE yordamida yorug'lik normalizatsiyasi; (2) Ko'p kanalli xususiyat ajratish moduli — HOG, LBP va CNN parallel kanallari; (3) Gibril klassifikator moduli — og'irlikli ansambl. Har bir modul alohida vazifani bajaradi va umumiy tizimning mustahkamligini oshiradi.

3.2. Yorug'lik normalizatsiyasi

Retinex nazariyasiga asosan tasvir ikkita komponentdan iborat: $S(x,y) = R(x,y) \cdot L(x,y)$, bu yerda S — kuzatilgan tasvir, R — reflektans, L — yoritilganlik komponenti. Yoritilganlikni ajratib olish:

$$R(x,y) = \log[S(x,y)] - \log[F(x,y) * S(x,y)] \quad (1)$$

bu yerda $F(x,y)$ — Gauss filtri, $*$ — konvolyutsiya operatori. So'ngra CLAHE qo'llaniladi:

$$T(r) = (L-1)/n \cdot \sum p(rk), \quad k = 0, 1, \dots, r \quad (2)$$

bu yerda L — gradatsiyalar soni, n — umumiy piksellar soni, $p(rk)$ — ehtimollik zichligi.

3.3. Ko'p kanalli xususiyat ajratish

HOG kanalida gradient yo'nalishlari hisoblanib, 9 ta yo'nalish bo'yicha gistogramma tuziladi. LBP kanalida har bir piksel p uchun:

$$LBP(p) = \sum s(g_n - g_c) \cdot 2^n, \quad n = 0, \dots, 7 \quad (3)$$

bu yerda g_c — markaziy piksel qiymati, g_n — qo'shni piksellar, $s(x) = 1$ agar $x \geq 0$, aks holda 0. CNN kanalida MobileNetV3 arxitekturasini qo'llaniladi — ImageNet asosida oldindan o'rgatilgan model yuz ma'lumotlarida qayta sozlanadi.

3.4. Gibril klassifikator

Uch kanaldan olingan xususiyat vektorlari og'irlikli ansambl usuli bilan birlashtiriladi:

$$D(x) = w_1 \cdot HOG_SVM(x) + w_2 \cdot LBP_SVM(x) + w_3 \cdot CNN(x) \quad (4)$$



bu yerda $w_1 = 0.3$, $w_2 = 0.3$, $w_3 = 0.4$ — og'irlik koeffitsientlari. CNN xususiyatlarining yuqori og'irligi uning semantik xususiyatlarni yaxshiroq ifodalashini aks ettiradi.

MUHOKAMA VA NATIJALAR

4.1. Ma'lumotlar to'plami va tajriba sharoiti

Tajribalar uchun uchta ochiq ma'lumotlar to'plamidan foydalanildi: (1) LFW (Labeled Faces in the Wild) — 13 233 tasvir, 5 749 shaxs; (2) Extended Yale B — 2 414 tasvir, 38 shaxs, turli yorug'lik sharoitlari; (3) AR Face Dataset — 4 000 tasvir, 126 shaxs. Qo'shimcha ravishda laboratoriya sharoitida 20 ishtirokchidan 3 600 tasvir yig'ildi (6 xil yorug'lik darajasi, 4 xil fon turi). Tajribalar Intel Core i7-11800H va NVIDIA RTX 3060 yordamida amalga oshirildi.

4.2. Taqqoslash natijalari

1-jadvalda taklif etilgan gibrid modelning turli usullar bilan taqqoslash natijalari keltirilgan.

1-jadval. Turli usullarning LFW ma'lumotlar to'plamidagi ishlash ko'rsatkichlari

Usul	Aniqlik (%)	F1-ball	Tezlik (FPS)	Past yorug'lik (%)
HOG + SVM	82.3	0.81	45	71.4
LBP + SVM	79.6	0.78	52	68.9
MobileNetV3 (CNN)	91.4	0.90	28	81.7
FaceNet	92.8	0.92	19	83.2
Gibrid model (taklif)	94.7	0.94	31	89.2

Jadvaldan ko'rinib turibdiki, taklif etilgan gibrid model barcha ko'rsatkichlar bo'yicha ustunlik qiladi. FaceNet bilan solishtirganda aniqlik 1.9% yuqori, tezlik esa 63% ortiq (31 FPS vs 19 FPS).

4.3. Yorug'lik sharoitiga qarab tahlil

2-jadvalda Extended Yale B to'plamida turli yorug'lik darajalarida erishilgan aniqlik ko'rsatkichlari keltirilgan.

2-jadval. Turli yorug'lik sharoitlarida aniqlik tahlili (Extended Yale B)

Yorug'lik darajasi	HOG+SVM (%)	CNN (%)	Gibrid model (%)
Yuqori (>200 lux)	88.1	94.3	96.8
O'rta (50–200 lux)	83.4	91.7	94.5
Past (10–50 lux)	71.4	81.7	89.2



Juda past (<10 lux)	58.2	69.4	76.1
---------------------	------	------	------

Yorug'lik qanchalik past bo'lsa, gibrid modelning ustunligi shunchalik yaqqol ko'zga tashlanadi. Past yorug'likda gibrid model CNN ga nisbatan 7.5% ko'proq aniqlik ko'rsatadi, bu Retinex normalizatsiyasining samaradorligini isbotlaydi.

4.4. Natijalar muhokamasi

Gibrid modelning muvaffaqiyati bir necha omil bilan izohlanadi. Birinchidan, Retinex asosidagi yorug'lik normalizatsiyasi tizimning turli yoritilganlik darajalariga moslashuvchanligini oshirdi. Ikkinchidan, ko'p kanalli xususiyat ajratish HOG (tuzilmaviy), LBP (to'qima) va CNN (semantik) xususiyatlarni bir vaqtda qamrab oladi. Uchinchidan, (4)-formuladagi og'irlikli ansambl klassifikatsiyasi alohida modullarning kuchli tomonlarini birlashtirishga imkon beradi. Asosiy cheklov — CNN modulining o'quv ma'lumotlariga bog'liqligi va yangi shaxslarni qo'shish uchun qayta o'qitish zarurligi.

XULOSA VA TAKLIFLAR

Ushbu tadqiqotda o'zgaruvchan yorug'lik va fon sharoitlarida yuzni aniqlash va tanib olish uchun yangi gibrid model taklif etildi. LFW ma'lumotlar to'plamida 94.7% aniqlik va 31 FPS tezligiga erishildi. Quyidagi asosiy xulosalar chiqarildi:

1. Ko'p kanalli xususiyat ajratish yondashuvi yuzni tanib olishning mustahkamligini sezilarli oshiradi — HOG, LBP va CNN xususiyatlarining to'ldiruvchanligi umumiy aniqlikni yaxshilaydi.

2. Retinex asosidagi yorug'lik normalizatsiyasi past yoritilganlik sharoitida aniqlikni 7.5% ga oshiradi.

3. Og'irlikli ansambl klassifikatsiyasi (4-formula) alohida usullar bilan solishtirganda muvozanatli ishlash ko'rsatkichlarini ta'minlaydi.

4. Taklif etilgan tizim real vaqt rejimida (31 FPS) xavfsizlik, tibbiyot va mobil ilovalar uchun qo'llanilishi mumkin.

Kelajakdagi tadqiqotlarda Vision Transformer (ViT) arxitekturasi va few-shot o'rganish usullarini integratsiya qilish, shuningdek, 3D yuz modellari yordamida tizim mustahkamligini yanada oshirish mo'ljallangan.

FOYDALANILGAN ADABIYOTLAR RO'YXATI:

[1] Zhao W., Chellappa R., Phillips P.J., Rosenfeld A. Face recognition: A literature survey // ACM Computing Surveys. — 2003. — Vol. 35, No. 4. — P. 399–458. DOI: 10.1145/954339.954342

[2] Kortli D., Tjahjadi T. A survey of face recognition applications // Pattern Recognition Letters. — 2020. — Vol. 131. — P. 232–240. DOI: 10.1016/j.patrec.2019.12.016





- [3] Wang M., Deng W. Deep face recognition: A survey // *Neurocomputing*. — 2021. — Vol. 429. — P. 215–244. DOI: 10.1016/j.neucom.2020.10.081
- [4] LeCun Y., Bengio Y., Hinton G. Deep learning // *Nature*. — 2015. — Vol. 521. — P. 436–444. DOI: 10.1038/nature14539
- [5] Viola P., Jones M. Rapid object detection using a boosted cascade of simple features // *Proc. IEEE CVPR*. — 2001. — Vol. 1. — P. 511–518. DOI: 10.1109/CVPR.2001.990517
- [6] Dalal N., Triggs B. Histograms of oriented gradients for human detection // *Proc. IEEE CVPR*. — 2005. — Vol. 1. — P. 886–893. DOI: 10.1109/CVPR.2005.177
- [7] Ojala T., Pietikäinen M., Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with LBP // *IEEE Trans. Pattern Anal. Mach. Intell.* — 2002. — Vol. 24, No. 7. — P. 971–987. DOI: 10.1109/TPAMI.2002.1017623
- [8] Taigman Y., Yang M., Ranzato M., Wolf L. DeepFace: Closing the gap to human-level performance in face verification // *Proc. IEEE CVPR*. — 2014. — P. 1701–1708. DOI: 10.1109/CVPR.2014.220
- [9] Schroff F., Kalenichenko D., Philbin J. FaceNet: A unified embedding for face recognition and clustering // *Proc. IEEE CVPR*. — 2015. — P. 815–823. DOI: 10.1109/CVPR.2015.7298682
- [10] Wang H. et al. Additive margin softmax for face verification // *IEEE Signal Process. Lett.* — 2018. — Vol. 25, No. 7. — P. 926–930. DOI: 10.1109/LSP.2018.2831071
- [11] Zhang Z., Luo P., Loy C.C., Tang X. Facial landmark detection by deep multi-task learning // *Proc. ECCV*. — 2014. — P. 94–108. DOI: 10.1007/978-3-319-10599-4_7
- [12] Ranjan R., Patel V.M., Chellappa R. Hyperface: A deep multi-task learning framework // *IEEE Trans. Pattern Anal. Mach. Intell.* — 2019. — Vol. 41, No. 1. — P. 121–135. DOI: 10.1109/TPAMI.2017.2781233

