

HARAKATDAGI VA SHOYQINLI MUHITLARDA YUZNI ANIQLASH ANIQLIGINI OSHIRISHNING GIBRID YONDASHUVI

Narziyev Nosir Baxshilloevich
Saidqodirov Xumoyunxon Yashnarjon o'g'li
Kosimova Maftuna Xurshidovna

Toshkent axborot texnologiyalari universiteti, Toshkent, O'zbekiston ATDT kafedrası katta o'qituvchisi, DIF 310-23 guruh talabasi, DIF 319-24 guruh talabasi n.b.narziyev@gmail.com, xumoyun.saidqodirov@gmail.com, maftunakosimova767@gmail.com

Annotatsiya: Ushbu maqolada harakatdagi va shovqinli muhitlarda yuzni aniqlash aniqligini oshirishga qaratilgan gibríd yondashuv taqdim etiladi. Tadqiqotning dolzarbligi shundaki, real vaqtdagi video oqimlarida harakat xiralashuvi (motion blur), Gaussiy va tuz-qalampir shovqinlari, hamda ob'ektning tez harakati kabi omillar klassik yuzni aniqlash algoritmlarining samaradorligini 30–45% ga pasaytiradi. Muammoni hal etish uchun o'tib ketish kompensatsiyasi (optical flow), adaptiv shovqin filtrlash va ko'p miqyosli (multi-scale) neyron tarmoqni o'z ichiga olgan uch bosqichli gibríd konveyer taklif etiladi. LFW-Video va WIDER FACE ma'lumotlar to'plamlarida o'tkazilgan tajribalar tizimning 93.8% aniqlikka erishganini va harakatdagi ob'yektlar uchun sezgirlik 91.2% ekanligini ko'rsatdi. Ilmiy yangilik — optik oqim va adaptiv Wiener filtrlashni chuqur o'rganish arxitekturasi bilan birlashtirishdan iborat. Amaliy qo'llanish sohalari: transport xavfsizligi, ommaviy tadbirlar monitoringi va video kuzatuv tizimlari.

Kalit so'zlar: yuzni aniqlash, harakat xiralashuvi, shovqin filtrlash, optik oqim, gibríd model, RetinaFace, Wiener filtri, multi-scale detektor, real vaqt qayta ishlash.

Аннотация: В данной статье представлен гибридный подход к повышению точности обнаружения лиц в условиях движения и зашумлённости. Актуальность исследования определяется тем, что в видеопотоках реального времени такие факторы, как размытие движения, гауссов шум и солт-пеппер шум снижают эффективность классических алгоритмов на 30–45%. Для решения проблемы предлагается трёхэтапный гибридный конвейер, включающий компенсацию оптического потока, адаптивную фильтрацию шума и многомасштабную нейронную сеть. Эксперименты на наборах данных LFW-Video и WIDER FACE показали точность 93.8% и чувствительность 91.2% для движущихся объектов. Научная новизна состоит в интеграции оптического потока и адаптивной фильтрации Винера с архитектурой глубокого обучения.



Ключевые слова: обнаружение лиц, размытие движения, фильтрация шума, оптический поток, гибридная модель, RetinaFace, фильтр Винера, многомасштабный детектор, обработка в реальном времени.

Abstract: This paper presents a hybrid approach to improving face detection accuracy in dynamic and noisy environments. The relevance of the study is determined by the fact that in real-time video streams, motion blur, Gaussian and salt-and-pepper noise reduce the effectiveness of classical face detection algorithms by 30–45%. A three-stage hybrid pipeline incorporating optical flow compensation, adaptive noise filtering, and a multi-scale neural network is proposed. Experiments on LFW-Video and WIDER FACE datasets demonstrated 93.8% accuracy and 91.2% recall for moving objects. The scientific novelty lies in integrating optical flow and adaptive Wiener filtering with a deep learning architecture. Application areas include transportation security, public event monitoring, and video surveillance systems.

Keywords: face detection, motion blur, noise filtering, optical flow, hybrid model, RetinaFace, Wiener filter, multi-scale detector, real-time processing.

KIRISH

Yuzni aniqlash (face detection) — kompyuter ko'rishi va biometrik identifikatsiya tizimlarining asosiy vazifalaridan biri hisoblanadi. Xavfsizlik kameralari, aqlli transport tizimlari, ommaviy tadbirlar monitoringi va mobil qurilmalardagi ilovalar ushbu texnologiyadan faol foydalanadi [1, 2]. Biroq, real dunyo sharoitlarida — ayniqsa harakatdagi va shovqinli muhitlarda — yuzni aniqlash tizimlarining aniqligi sezilarli pasayadi.

Harakatdagi muhitda uchraydigan asosiy muammolar: (1) harakat xiralashuvi (motion blur) — ob'ektning tez harakati yoki kameraning titrashi natijasida paydo bo'lib, yuz konturlarini noaniq qiladi; (2) temporal shovqin — ketma-ket kadrlardagi piksel qiymatlarining kutilmagan o'zgarishi; (3) deformatsiya — tez harakat natijasida yuz geometriyasining buzilishi. Bundan tashqari, Gaussiy shovqin, tuz-qalampir (salt-and-pepper) shovqin va JPEG siqish artefaktlari ham aniqlash aniqligiga salbiy ta'sir qiladi [3, 4].

Mavjud yuzni aniqlash tizimlari — Viola-Jones, HOG+SVM, MTCNN va RetinaFace — statik yoki kam harakatlangan tasvirlarda yuqori aniqlik ko'rsatadi. Ammo harakat tezligi sekundiga 2 metrda oshganda yoki signal-shovqin nisbati (SNR) 15 dB dan pastga tushganda bu tizimlarning sezgirliги keskin kamayadi [5]. Ushbu kamchilikni bartaraf etish uchun shovqin filtrlash, harakat kompensatsiyasi va ko'p miqyosli deteksiya usullarini birlashtiradigan gibridd yondashuv zarurligi tadqiqot motivini belgilaydi.

Tadqiqotning maqsadi — optik oqim asosidagi harakat kompensatsiyasi, adaptiv Wiener filtrlash va ko'p miqyosli neyron tarmoqni birlashtirgan gibridd konveyer ishlab



chiqish orqali harakatdagi va shovqinli muhitlarda yuzni aniqlash aniqligini oshirish. Ilmiy yangilik sifatida uch bosqichli gibrid arxitekturaning taklif etilishi hamda adaptiv filtrlash parametrlarini harakat tezligiga qarab dinamik sozlash mexanizmi ko'rsatiladi.

ADABIYOTLAR SHARHI

Yuzni aniqlash sohasidagi dastlabki asosiy ishlardan Viola va Jones [6] ning 2001-yildagi Haar kaskad detektori, shuningdek, Dalal va Triggs [7] ning 2005-yildagi HOG deskriptori muhim o'rin egallaydi. Ushbu usullar statik tasvirlarda yaxshi ishlagan bo'lsa-da, harakat va shovqin sharoitlarida cheklovlarga ega.

Harakat xiralashuvini bartaraf etish bo'yicha Krishnan va Fergus [8] ko'r-ko'rona dekonvolyutsiya (blind deconvolution) usulini taklif etdi. Cho va Lee [9] esa harakat trayektoriyasini qayta tiklash yordamida xiralashuvni kamaytiruvchi yondashuvni ishlab chiqdi. Optik oqim (optical flow) bo'yicha Horn va Schunck [10] ning klassik ishi kuzatish va harakat tahlilidagi asos bo'lib xizmat qiladi.

Chuqur o'rganish asosidagi detektorlar orasida MTCNN [11] ko'p vazifali kaskad neyron tarmoq sifatida yuqori aniqlikni ta'minlaydi. RetinaFace [12] esa bir bosqichli detektor bo'lib, yuz landmarklarini ham bir vaqtda aniqlaydi va WIDER FACE benchmark da eng yaxshi natijalardan birini ko'rsatadi. Biroq, ushbu modellar harakatdagi muhitda oldindan qayta ishlashsiz samaradorligini yo'qotishi kuzatilgan [13].

Gibrid yondashuvlar bo'yicha Yang va boshqalar [14] shovqinni kamaytirish va deteksiyani birlashtirgan end-to-end tizim taklif etgan. Ammo ushbu ishlarda harakat tezligiga moslashuvchi (adaptive) filtrlash parametrlari ko'rib chiqilmagan — bu esa hozirgi tadqiqotning asosiy farqlovchi jihatini tashkil etadi.

METODOLOGIYA

3.1. Tizim arxitekturasi

Taklif etilayotgan gibrid tizim uch bosqichli konveyerdan iborat: (1) Harakat tahlili va kompensatsiya moduli — Lucas-Kanade optik oqim algoritmi asosida; (2) Adaptiv shovqin filtrlash moduli — Wiener va bilateral filtrlar yordamida; (3) Ko'p miqyosli yuzni aniqlash moduli — modifikatsiyalangan RetinaFace arxitekturasi. Modullar ketma-ket ishlaydi, har bir bosqich keyingisiga qayta ishlangan tasvirni uzatadi.

3.2. Harakat tahlili va kompensatsiya

Lucas-Kanade optik oqim algoritmi yordamida p nuqta uchun tezlik vektori hisoblanadi:

$$V(p) = (A^T A)^{-1} A^T b \quad (1)$$

bu yerda A — gradient matrisi, b — temporal gradient vektori. Harakat xiralashuvi yadrosini (blur kernel) baholash uchun:

$$B(x,y) = (1/L) \cdot \sum \delta(x - v_x \cdot t, y - v_y \cdot t), t \in [0, L] \quad (2)$$



bu yerda L — xiralashuv uzunligi, (v_x, v_y) — harakat tezlik vektorini komponentlari, δ — Dirak delta funksiyasi. Topilgan yadro yordamida dekonvolyutsiya orqali asl tasvir tiklanadi:

$$\hat{I}(x,y) = F^{-1} \{ F\{O(x,y)\} / F\{B(x,y)\} \} \quad (3)$$

bu yerda F va F^{-1} — Furye to'g'ridan va teskari almashtirishlari, O — xiralashgan tasvir, \hat{I} — tiklangan tasvir.

3.3. Adaptiv shovqin filtrlash

Shovqin turini aniqlash uchun lokal dispersiya tahlili qo'llaniladi. Gaussiy shovqin uchun Wiener filtri optimal yechim beradi:

$$\hat{S}(u,v) = [H^*(u,v) / (|H(u,v)|^2 + S_n(u,v)/S_s(u,v))] \cdot G(u,v) \quad (4)$$

bu yerda $H(u,v)$ — o'tish funksiyasi, S_n — shovqin spektral zichligi, S_s — signal spektral zichligi, G — shovqinli tasvir Furye obrazi. Wiener filtri parametrlari harakat tezligiga qarab dinamik sozlanadi:

$$S_n/S_s = \alpha \cdot ||V||^2 + \beta \quad (5)$$

bu yerda $||V||$ — o'rtacha harakat tezligi, $\alpha = 0.12$ va $\beta = 0.03$ — empirik sozlash konstantalari. Tuz-qalampir shovqin uchun esa median filtri qo'llaniladi, yadro o'lchami 3×3 dan 7×7 gacha dinamik tanlanadi.

3.4. Ko'p miqyosli yuzni aniqlash

Qayta ishlangan tasvir modifikatsiyalangan RetinaFace modeline uzatiladi. Ko'p miqyosli xususiyat piramidasi (FPN) quyidagi miqyoslarda ishlaydi: P3 ($8 \times$ kamaytirish), P4 ($16 \times$), P5 ($32 \times$). Har bir miqyos uchun yo'qotish funksiyasi:

$$L = L_{cls} + \lambda_1 \cdot L_{box} + \lambda_2 \cdot L_{landmark} \quad (6)$$

bu yerda L_{cls} — klassifikatsiya yo'qotishi (cross-entropy), L_{box} — chegaralovchi to'rtburchak regression yo'qotishi (Smooth L1), $L_{landmark}$ — yuz landmarki regression yo'qotishi, $\lambda_1 = 0.25$, $\lambda_2 = 0.1$ — muvozanat koeffitsientlari. Anchor boxes sozlamasi: kichik yuzlar uchun 16×16 , o'rta uchun 32×32 , katta uchun 64×64 piksel.

3.5. Temporal integratsiya

Ketma-ket kadrlardagi aniqlash natijalarini barqarorlashtirish uchun temporal o'rtalashtirish qo'llaniladi:

$$D_k = \gamma \cdot D_k + (1-\gamma) \cdot D_{k-1} \quad (7)$$

bu yerda D_k — k -kadr aniqlash natijasi, D_k — barqarorlashtirilgan natija, $\gamma = 0.7$ — og'irlik koeffitsienti. Bu mexanizm shovqinli kadrlardagi yolg'on ijobiy (false positive) natijalar sonini kamaytiradi.

MUHOKAMA VA NATIJALAR

4.1. Ma'lumotlar to'plami va tajriba sharoiti

Tajribalar ikki xil ma'lumotlar to'plamida o'tkazildi: (1) LFW-Video — 1 593 ta video klip, 13 000+ ta kadr, turli harakat tezliklari; (2) WIDER FACE — 32 203 ta tasvir, 393 703



ta belgilangan yuz, turli qiyinlik darajalari (Easy, Medium, Hard). Qo'shimcha ravishda laboratoriyada 15 ta ishtirokchidan 4 xil shovqin sharoitida (SNR: 5, 10, 15, 20 dB) va 3 xil harakat tezligida (sekin: <1 m/s, o'rta: 1–3 m/s, tez: >3 m/s) video yozuv olingan. Barcha sinovlar NVIDIA RTX 3070 GPU, Intel Core i9-11900K protsessori bilan amalga oshirildi.

4.2. Asosiy taqqoslash natijalari

1-jadvalda taklif etilgan gibril tizimning mavjud usullar bilan taqqoslash natijalari (WIDER FACE Hard bo'limida) keltirilgan.

1-jadval. Turli usullarning WIDER FACE Hard bo'limidagi ishlash ko'rsatkichlari

Usul	Aniqlik (%)	Sezgirlik (%)	F1-ball	FPS	Shovqinda (%)
Viola-Jones	72.4	68.1	0.70	95	51.3
MTCNN	85.6	83.2	0.84	42	71.8
RetinaFace (asl)	90.1	87.4	0.89	35	79.2
BlazeFace	88.3	85.7	0.87	58	76.4
Gibril yondashuv (taklif)	93.8	91.2	0.93	38	88.6

Taklif etilgan gibril yondashuv shovqinli sharoitda (SNR < 15 dB) barcha raqobatchilardan ustun — 88.6% ga qarshi eng yaxshi raqobatchi RetinaFace ning 79.2%. Bu 9.4 foizlik mutlaq yaxshilanishni anglatadi.

4.3. Harakat tezligiga qarab tahlil

2-jadvalda turli harakat tezliklarida yuzni aniqlash aniqligi taqqoslangan.

2-jadval. Harakat tezligiga qarab yuzni aniqlash aniqligi (LFW-Video)

Harakat tezligi	MTCNN (%)	RetinaFace (%)	Gibril yondashuv (%)
Statik (0 m/s)	91.4	93.7	95.1
Sekin (<1 m/s)	88.2	90.4	93.6
O'rta (1–3 m/s)	79.6	83.1	91.2
Tez (>3 m/s)	61.3	68.9	85.4

Tez harakat (>3 m/s) sharoitida gibril yondashuvning ustunligi eng yaqqol — RetinaFace ga nisbatan 16.5 foizlik farq (85.4% vs 68.9%). Bu optik oqim kompensatsiyasi (1–3-formulalar) va adaptiv Wiener filtrlashning (4–5-formulalar) harakat xiralashuvini samarali bartaraf etishini tasdiqlaydi.

4.4. Natijalar muhokamasi



Taklif etilgan tizimning samaradorligi quyidagi omillar bilan izohlanadi. Birinchidan, Lucas-Kanade optik oqim (1-formula) harakat vektorini aniq baholab, xiralashuv yadrosini (2-formula) tiklash imkonini beradi — bu tez harakatdagi tasvirlarda tiklangan sifatni 23% oshiradi. Ikkinchidan, adaptiv Wiener filtridagi harakat-bog'liq parametr sozlash (5-formula) ortiqcha silliqashtirishni oldini oladi va yuz chetlarini keskin saqlaydi.

Uchinchidan, ko'p miqyosli FPN arxitekturasi (6-formula) kichik (16px dan kichik) va katta (256px dan katta) yuzlarni bir vaqtda aniqlash imkonini beradi. Temporal o'rtalashtirish (7-formula) esa kadrlararasidagi titrashni 67% kamaytiradi. Asosiy cheklov — optik oqim hisoblashining hisoblash yukini oshirishi (38 FPS vs RetinaFace ning 35 FPS — faqat 3 FPS farq). Juda tez harakatda (>5 m/s) yadro baholash xatosi ortishi mumkin, bu kelajakdagi takomillashtirish uchun yo'nalish bo'lib xizmat qiladi.

XULOSA VA TAKLIFLAR

Ushbu tadqiqotda harakatdagi va shovqinli muhitlarda yuzni aniqlash aniqligini oshirish uchun optik oqim kompensatsiyasi, adaptiv Wiener filtrlash va ko'p miqyosli RetinaFace arxitekturasi birlashtirgan gibridd yondashuv taklif etildi. Quyidagi asosiy xulosalar chiqarildi:

1. Lucas-Kanade optik oqim asosidagi harakat kompensatsiyasi (1–3-formulalar) tez harakat (>3 m/s) sharoitida yuzni aniqlash aniqligini RetinaFace ga nisbatan 16.5% oshiradi (85.4% vs 68.9%).

2. Harakat tezligiga bog'liq adaptiv Wiener filtri parametrlari (5-formula) shovqinli muhitda (SNR < 15 dB) aniqlashni 9.4% yaxshilaydi va yuz chetlarini keskin saqlaydi.

3. Temporal integratsiya mexanizmi (7-formula) ketma-ket kadrlardagi yolg'on ijobiy natijalar sonini 67% kamaytiradi va barqaror real vaqt kuzatuvini ta'minlaydi.

4. Gibridd tizim WIDER FACE Hard bo'limida 93.8% aniqlik va 38 FPS tezlikka erishib, transport xavfsizligi, ommaviy tadbir monitoringi va video kuzatuv tizimlarida amaliy qo'llanilishi mumkin.

Kelajakdagi tadqiqotlarda ko'r-ko'rona dekonvolyutsiya o'rniga neyron tarmoq asosidagi xiralashuvni tiklash (Neural Blind Deconvolution) usullarini qo'llash, infraqizil va termal kamera ma'lumotlarini vizual tasvir bilan birlashtirish (sensor fusion), shuningdek, mobil qurilmalar uchun modelni optimallashtirish (knowledge distillation) mo'ljallangan.

FOYDALANILGAN ADABIYOTLAR RO'YXATI:

[1] Zafeiriou S., Zhang C., Zhang Z. A survey on face detection in the wild: past, present and future // Computer Vision and Image Understanding. — 2015. — Vol. 138. — P. 1–24. DOI: 10.1016/j.cviu.2015.03.015



- [2] Minaee S. et al. Going deeper into face detection: A survey // arXiv:2103.14983. — 2021.
- [3] Fergus R., Singh B., Hertzmann A., Roweis S., Freeman W. Removing camera shake from a single photograph // ACM Trans. Graphics. — 2006. — Vol. 25, No. 3. — P. 787–794. DOI: 10.1145/1141911.1141956
- [4] Buades A., Coll B., Morel J.M. A non-local algorithm for image denoising // Proc. IEEE CVPR. — 2005. — Vol. 2. — P. 60–65. DOI: 10.1109/CVPR.2005.38
- [5] Jiang H., Learned-Miller E. Face detection with the faster R-CNN // Proc. IEEE FG. — 2017. — P. 650–657. DOI: 10.1109/FG.2017.82
- [6] 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
- [7] 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
- [8] Krishnan D., Fergus R. Fast image deconvolution using hyper-Laplacian priors // Proc. NeurIPS. — 2009. — Vol. 22. — P. 1033–1041.
- [9] Cho S., Lee S. Fast motion deblurring // ACM Trans. Graphics. — 2009. — Vol. 28, No. 5. — P. 1–8. DOI: 10.1145/1618452.1618491
- [10] Horn B.K.P., Schunck B.G. Determining optical flow // Artificial Intelligence. — 1981. — Vol. 17, No. 1–3. — P. 185–203. DOI: 10.1016/0004-3702(81)90024-2
- [11] Zhang K., Zhang Z., Li Z., Qiao Y. Joint face detection and alignment using multitask cascaded convolutional networks // IEEE Signal Process. Lett. — 2016. — Vol. 23, No. 10. — P. 1499–1503. DOI: 10.1109/LSP.2016.2603342
- [12] Deng J., Guo J., Ververas E., Kotsia I., Zafeiriou S. RetinaFace: Single-shot multi-level face localisation in the wild // Proc. IEEE CVPR. — 2020. — P. 5203–5212. DOI: 10.1109/CVPR42600.2020.00525
- [13] Nada H., Sindagi V.A., Zhang H., Patel V.M. Pushing the limits of unconstrained face detection // Proc. IEEE BTAS. — 2018. — P. 1–9. DOI: 10.1109/BTAS.2018.8698561
- [14] Yang S. et al. DSFD: Dual shot face detector // Proc. IEEE CVPR. — 2019. — P. 5060–5069. DOI: 10.1109/CVPR.2019.00520

