DOI: 10.65322/j.srd.05192026.128

YIRIK VIZUAL MODELDAN BILIM DISTILLATSIYASI ORQALI KO‘Z KUZATUVI TIZIMINI ISHLAB CHIQISH

Safarov F.E., *dotsent v.b.,*
e-mail: sfurkat091@gmail.com
Xolmurodov S.L., *Kompyuter injiniringi mutaxassisligi magistranti,*
e-mail: sarvarbekxolmurodov39@gmail.com
Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti

Annotatsiya. *Ushbu maqolada yirik vizual modeldan bilim distillatsiyasi (knowledge distillation) orqali ko‘z tasviridan qarash yo‘nalishini baholovchi, qurilmada ishlashga mo‘ljallangan yengil modelni ishlab chiqish masalasi ko‘rib chiqilgan. Muzlatilgan DINOv2 ViT-S/14 kodlovchisi va Ridge regressiyasidan iborat o‘qituvchi model hamda 146 786 parametrli konvolyutsion neyron tarmoq ko‘rinishidagi o‘quvchi model qurildi. Tajribalar MPIIGaze to‘plamining 15 ishtirokchisiga tegishli 1500 ta haqiqiy ko‘z tasvirida, ma’lumotlarni ishtirokchilar bo‘yicha ajratish sharoitida o‘tkazildi. Bilim distillatsiyasi bilan o‘qitilgan model mustaqil test to‘plamida 7,05° o‘rtacha burchak xatosiga erishdi, bu oddiy usulda o‘qitilgan modeldan 4,56 % ga, o‘qituvchi modeldan esa 15,8 % ga kam. O‘quvchi model o‘qituvchidan 150,3 marta kichik va CPU da taxminan 39 marta tez ishlaydi. Natijalar bitta tasodifiy urug‘ va ikki test ishtirokchisida olingani sababli dastlabki baho sifatida qaralishi lozim.*

Аннотация. *В статье рассматривается задача разработки облегчённой модели оценки направления взгляда по изображению глаза, предназначенной для работы непосредственно на устройстве, с использованием дистилляции знаний (knowledge distillation) из крупной визуальной модели. Построены модель-учитель на основе замороженного кодировщика DINOv2 ViT-S/14 и гребневой регрессии, а также модель-ученик в виде свёрточной нейронной сети со 146 786 параметрами. Эксперименты проведены на 1500 реальных изображениях глаз 15 участников набора данных MPIIGaze при разделении данных по участникам. Модель, обученная с дистилляцией знаний, достигла средней угловой ошибки 7,05° на независимой тестовой выборке, что на 4,56 % меньше, чем у модели с обычным обучением, и на 15,8 % меньше, чем у модели-учителя. Модель-ученик в 150,3 раза меньше учителя и работает на CPU примерно в 39 раз быстрее. Поскольку результаты получены при одном случайном зерне и двух тестовых участниках, их следует рассматривать как предварительные.*

Abstract. *This article addresses the development of a lightweight on-device model that estimates gaze direction from an eye image using knowledge distillation from a large vision foundation model. A teacher model based on a frozen DINOv2 ViT-S/14 encoder with a ridge regression head and a convolutional neural network student with 146,786 parameters were built. Experiments were conducted on 1,500 real eye images from 15 participants of the MPIIGaze dataset using a participant-wise data split. The student trained with knowledge distillation achieved a mean angular error of 7.05° on the held-out test set, which is 4.56% lower than that of the conventionally trained student and 15.8% lower than that of the teacher. The student is 150.3 times smaller than the teacher and runs about 39 times faster on a CPU. Since the results were obtained with a single random seed and two test participants, they should be regarded as preliminary.*

Kalit so‘zlar: *qarash yo‘nalishini baholash, ko‘z kuzatuvi, bilim distillatsiyasi, DINOv2, konvolyutsion neyron tarmoq, MPIIGaze, qurilmada hisoblash.*

Ключевые слова: *оценка направления взгляда, отслеживание взгляда, дистилляция знаний, DINOv2, свёрточная нейронная сеть, MPIIGaze, вычисления на устройстве.*

DOI: 10.65322/j.srd.05192026.128

Keywords: *gaze estimation, eye tracking, knowledge distillation, DINOv2, convolutional neural network, MPIIGaze, on-device inference.*

Kirish. Ko‘z qarash yo‘nalishini aniqlash (gaze estimation) inson–kompyuter interfeyslari, kengaytirilgan va virtual reallik qurilmalari hamda haydovchi diqqatini nazorat qilish tizimlarining muhim tarkibiy qismi hisoblanadi [1]. Bunday tizimlar odatda smartfon, avtomobil bortidagi kompyuter yoki garnitura kabi hisoblash resurslari cheklangan qurilmalarda real vaqt rejimida ishlashi talab etiladi. Shu sababli qarash yo‘nalishini baholovchi model bir vaqtning o‘zida ham aniq, ham yengil bo‘lishi lozim.

2015-yilda X. Zhang va hammualliflar kundalik sharoitda noutbuk kamerasi orqali yig‘ilgan MPIIGaze to‘plamini taqdim etib, qarash yo‘nalishini ko‘z tasviridan bevosita konvolyutsion neyron tarmoq yordamida baholash mumkinligini ko‘rsatdilar [2, 3]. 2016-yilda K. Krafka va hammualliflar smartfon kamerasi tasvirlari asosida ishlovchi ko‘z kuzatuvi tizimini taklif etib, bu yo‘nalishni mobil qurilmalarga olib chiqdilar [4]. Shu bilan bir qatorda, G. Hinton va hammualliflar katta modelning bilimini kichik modelga o‘tkazish g‘oyasini – bilim distillatsiyasi usulini taklif qildilar [5].

2021–2024-yillarda vizual transformerlar [6] va DINOv2 kabi o‘z-o‘zini nazorat qilish (self-supervised) asosida o‘qitilgan yirik vizual asos modellar [7] paydo bo‘lib, ular turli vazifalarda yuqori umumlashtirish qobiliyatini namoyish etdi. 2026-yilda taklif etilgan DistillGaze yondashuvida esa vizual asos model bilimlari distillatsiya qilinib, qurilmada ishlovchi 256 ming parametrli ko‘z kuzatuvi modeli olingan [1].

Shu bilan birga, mavjud tadqiqotlar tahlili shuni ko‘rsatadiki, yirik vizual modellarning o‘nlab million parametrga ega ekani ularni qurilmaga bevosita joylashtirishni qiyinlashtiradi, distillatsiya bo‘yicha ishlar esa ko‘pincha yopiq va katta hajmli ma’lumotlarda bajarilgan. Ochiq to‘plamlarda, ma’lumotlarni ishtirokchilar bo‘yicha ajratish sharoitida yirik modeldan kichik modelga bilim distillatsiyasining samarasini aniqlik, hajm va tezlik bo‘yicha birgalikda baholash masalasi yetarlicha o‘rganilmagan.

Mualliflar fikricha, qurilmada ishlovchi ko‘z kuzatuvi tizimlarining asosiy muammosi yirik modelning aniqligi bilan kichik modelning tezligi o‘rtasidagi ziddiyatdir. Bu ziddiyatni hal qilishning samarali yo‘li yirik modelni bevosita ishlatish emas, balki undan o‘qituvchi (teacher) sifatida foydalanib, uning bilimini kichik o‘quvchi (student) modelga o‘tkazishdir. Ushbu ishning maqsadi DistillGaze g‘oyasi asosida ko‘z tasviridan qarash yo‘nalishining pitch va yaw burchaklarini baholovchi yengil modelni ishlab chiqish hamda uning aniqligi, hajmi va tezligini baholashdan iborat.

Tadqiqot obyekti va usullari. ***Masalaning qo‘yilishi.*** Kirish sifatida 36×60 piksel o‘lchamli kulrang (grayscale) ko‘z tasviri olinadi, chiqishda esa qarash yo‘nalishini ifodalovchi ikki burchak – pitch (θ, vertikal) va yaw (φ, gorizontal) radianlarda hisoblanadi. Shu sababli masala regressiya masalasi sifatida qo‘yiladi. Ishda faqat chap ko‘z tasvirlari qo‘llanildi. Burchaklar juftligi uch o‘lchamli birlik vektorga quyidagicha o‘tkaziladi:

$$g = (-\cos\theta \cdot \sin\varphi;\ -\sin\theta;\ -\cos\theta \cdot \cos\varphi). \quad (1)$$

Bilim distillatsiyasi. Bilim distillatsiyasida o‘quvchi model faqat haqiqiy belgilardan emas, balki o‘qituvchi model bashoratlaridan ham o‘rganadi [5]. Masala klassifikatsiya emas, regressiya bo‘lgani uchun softmax harorati qo‘llanilmaydi va umumiy yo‘qotish funksiyasi quyidagicha aniqlanadi:

$$L_{\mathrm{KD}} = (1-\alpha) \cdot \mathrm{MSE}(S(x), y) + \alpha \cdot \mathrm{MSE}(S(x), T(x)), \quad (2)$$

bu yerda $S(x)$ va $T(x)$ – o‘quvchi va o‘qituvchi model bashoratlari, y – haqiqiy burchaklar, $\alpha = 0{,}3$ – o‘qituvchi bashoratlarining ulushini belgilovchi koeffitsient. $\alpha = 0$ bo‘lganda (2) ifoda oddiy o‘qitishga (baseline) aylanadi. α qiymati ablatsiya tajribasisiz, dastlabki tanlov sifatida olingan.

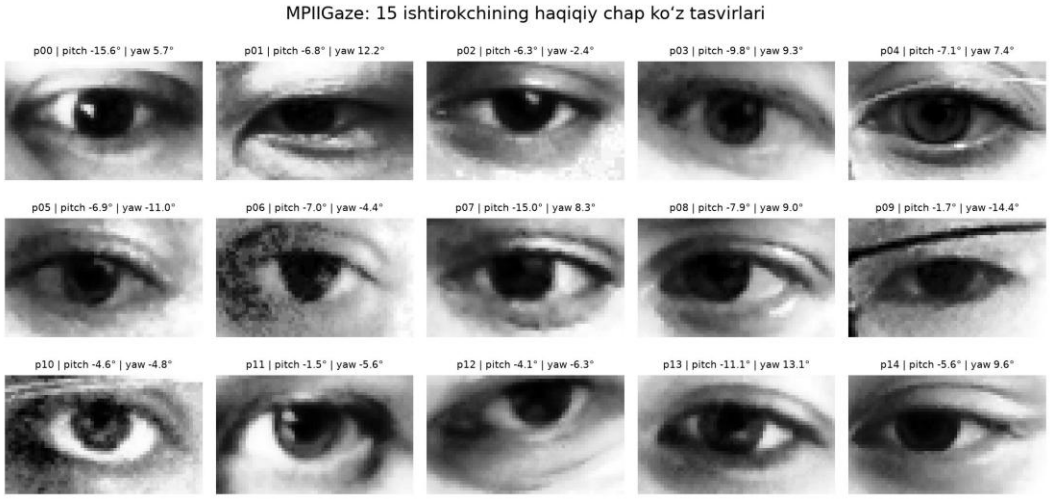
DOI: 10.65322/j.srd.05192026.128

Baholash mezoni. Modellar aniqligi haqiqiy va bashorat qilingan qarash vektorlari orasidagi burchak bilan baholanadi:

$$e = \arccos(g_{\text{pred}} \cdot g_{\text{true}}) \cdot \frac{180}{\pi}, \quad (3)$$

bu yerda g_{pred} va g_{true} – (1) ifoda bo‘yicha hisoblangan bashorat va haqiqiy birlik vektorlar. Qo‘shimcha ravishda median xato, 90-protsentil (P90) xato va xatosi 5° dan kichik bo‘lgan tasvirlar ulushi hisoblandi.

Ma’lumotlar to‘plami. Tajribada MPIIGaze ochiq ma’lumotlar to‘plami qo‘llanildi [3, 8]. U kundalik sharoitda noutbuk kamerasi orqali olingan va geometrik normalizatsiya qilingan ko‘z tasvirlaridan iborat. To‘plamning 15 ishtirokchisidan (p00–p14) har biridan 100 tadan, jami 1500 ta haqiqiy chap ko‘z tasviri tanlandi (1-rasm).



1-rasm. MPIIGaze to‘plamidagi ishtirokchilarning chap ko‘z tasvirlari va burchak belgilari

Modelning yangi odamlarga umumlashtirish qobiliyatini xolis baholash uchun ma’lumotlar ishtirokchilar bo‘yicha ajratildi: bir odamning tasvirlari bir vaqtning o‘zida o‘qitish va test qismlarida qatnashmaydi, modelga shaxs identifikatori berilmaydi. Validatsiya qismi faqat model va epoxani tanlash uchun, test qismi esa faqat yakuniy baholash uchun ishlatildi (1-jadval).

1-jadval

Ma’lumotlar to‘plamining ajratilishi

Qism	Ishtirokchilar	Tasvirlar soni	Vazifasi
O‘qitish	p00–p10	1100	Modellarni o‘qitish
Validatsiya	p11–p12	200	Model va epoxani tanlash
Test	p13–p14	200	Yakuniy baholash
Jami	15 nafar	1500	–

O‘qituvchi va o‘quvchi modellar. O‘qituvchi model sifatida muzlatilgan DINOv2 ViT-S/14 kodlovchisi qo‘llanildi [7], uning vaznlari o‘qitish jarayonida o‘zgartirilmaydi. Kulrang ko‘z tasviri uch kanalga takrorlanib, 70×112 piksel o‘lchamga keltiriladi, kodlovchining CLS belgisi va patch belgilarining o‘rtachasi birlashtirilib, 768 o‘lchamli xususiyat vektori hosil qilinadi. Xususiyatlar Ridge regressiyasiga [9] beriladi, regulyarizatsiya koeffitsienti $\lambda \in \{1; 10; 100; 1000\}$ qiymatlari orasidan validatsiya xatosi bo‘yicha tanlandi ($\lambda = 1000$). O‘qituvchi modelning umumiy parametrlar soni 22 058 114 ta.

O‘quvchi model 36×60 o‘lchamli kulrang tasvirdan ikki burchakni bashorat qiluvchi kichik konvolyutsion tarmoq bo‘lib, u uchta konvolyutsion blok va ikki qatlamli regressiya boshidan

DOI: 10.65322/j.srd.05192026.128

iborat (2-jadval). Fazoviy joylashuv ma'lumotini saqlash maqsadida oxirgi blokdan keyin 3×5 o'lchamli moslashuvchan o'rtacha pooling qo'llanildi.

2-jadval

O'quvchi model arxitekturasi

№	Qatlam	Sozlamalar	Chiqish o'lchami	Parametrlar soni
1	Conv2d, GroupNorm, SiLU	1→16, 3×3, stride 2	16×18×30	192
2	Conv2d, GroupNorm, SiLU	16→32, 3×3, stride 2	32×9×15	4 704
3	Conv2d, GroupNorm, SiLU	32→64, 3×3, stride 2	64×5×8	18 624
4	AdaptiveAvgPool2d, Flatten	3×5	960	0
5	Linear, SiLU, Dropout	960→128, p = 0,1	128	123 008
6	Linear	128→2	2	258
	Jami			**146 786**

Shunday qilib, o'quvchi model 146 786 ta parametrga ega bo'lib, o'qituvchi modeldan 150,3 marta kichik.

Algoritm va dastur. Taklif etilgan yondashuvni amaliyotda qo'llash maqsadida bilim distillatsiyasi asosida yengil model olish algoritmi ishlab chiqildi. Algoritm quyidagi bosqichlardan iborat:

1) *Ma'lumotlarni tayyorlash.* MPIIGaze rasmiy arxividan kerakli MAT fayllar yuklab olinadi. Chap ko'z tasvirlari va ularning [pitch, yaw] belgilari ajratib olinib, ishtirokchilar bo'yicha o'qitish, validatsiya va test qismlariga bo'linadi.

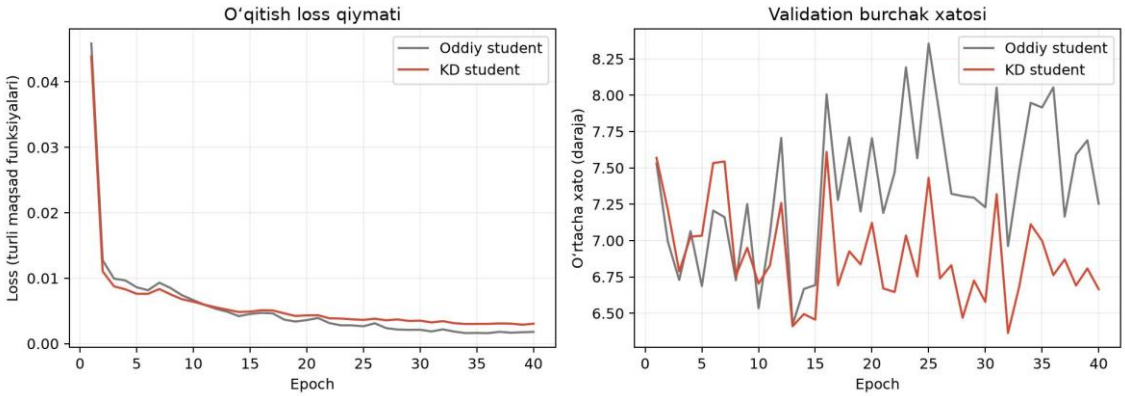
2) *O'qituvchi modelni qurish.* Muzlatilgan DINOv2 kodlovchisidan barcha tasvirlar uchun xususiyat vektorlari olinadi. StandardScaler faqat o'qitish qismi statistikasi bo'yicha moslanadi va Ridge regressiyasi o'qitiladi, so'ngra o'qituvchi bashoratlari T(x) barcha tasvirlar uchun hisoblanadi.

3) *O'quvchi modelni o'qitish.* Oddiy va KD o'quvchilar bir xil boshlang'ich vaznlar va tasodifiy urug' (seed = 42) bilan, AdamW optimallashtiruvchisi (o'rganish tezligi 0,001, batch o'lchami 64) yordamida 40 epoxa davomida (2) yo'qotish funksiyasi bo'yicha o'qitiladi. Har bir epoxa oxirida validatsiya xatosi hisoblanib, eng kichik xatoga ega epoxa vaznlari saqlab qolinadi.

4) *Baholash va eksport.* Yakuniy modellar test qismida (3) ifoda bo'yicha baholanadi, tezlik esa CPU da batch = 1 rejimida o'lchanadi. So'ngra KD o'quvchi modeli TorchScript formatiga eksport qilinib, qayta yuklangan model natijalari asl model natijalari bilan solishtiriladi.

Algoritm Python dasturlash tilida PyTorch kutubxonasi [10] yordamida dasturlashtirildi. O'qitish jarayonining dinamikasi 2-rasmda ko'rsatilgan: oddiy o'quvchi uchun eng yaxshi natija 13-epoxada (6,42°), KD o'quvchi uchun esa 32-epoxada (6,36°) olindi.

DOI: 10.65322/j.srd.05192026.128



2-rasm. O‘qitishdagi yo‘qotish funksiyasi (chapda) va validatsiya burchak xatosining (o‘ngda) epoxalar bo‘yicha o‘zgarishi

Natijalar va muhokama. Yakuniy baholash o‘qitish va model tanlashda qatnashmagan ikki ishtirokchining (p13, p14) 200 ta tasvirida amalga oshirildi. Modellarning aniqlik va tezlik ko‘rsatkichlari 3-jadvalda keltirilgan, har bir ko‘rsatkich bo‘yicha eng yaxshi qiymat qalin shriftda ajratilgan.

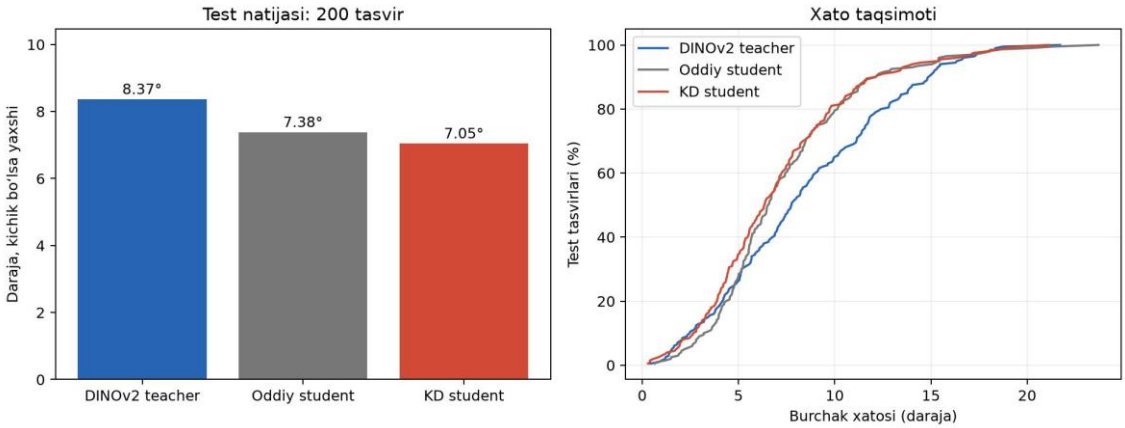
3-jadval

Modellarning test to‘plamidagi natijalari va hisoblash tezligi

Ko‘rsatkich	DINOv2 o‘qituvchi	Oddiy o‘quvchi	KD o‘quvchi
O‘rtacha burchak xatosi, °	8,37	7,38	**7,05**
Median xato, °	7,76	6,57	**6,36**
P90 xato, °	14,87	12,14	**12,10**
Xatosi 5° dan kichik tasvirlar, %	26,0	27,5	**34,5**
Pitch bo‘yicha o‘rtacha absolyut xato, °	**3,16**	4,23	3,75
Yaw bo‘yicha o‘rtacha absolyut xato, °	7,24	**5,23**	5,26
Parametrlar soni	22 058 114	146 786	146 786
Median kechikish (CPU), ms	10,11	**0,23**	0,26

Jadvaldan ko‘rinadiki, bilim distillatsiyasi bilan o‘qitilgan o‘quvchi model eng kichik o‘rtacha xatoga (7,05°) erishdi: bu oddiy o‘quvchidan 0,34° ga (4,56 %), o‘qituvchi modeldan esa 15,8 % ga kam. Xatosi 5° dan kichik tasvirlar ulushi 27,5 % dan 34,5 % gacha oshdi. Test xatolarining taqsimoti 3-rasmda keltirilgan.

DOI: 10.65322/j.srd.05192026.128



3-rasm. Test to‘plamidagi o‘rtacha burchak xatosi (chapda) va xatolarning kumulyativ taqsimoti (o‘ngda)

Kumulyativ taqsimot grafigida o‘quvchi modellarning egri chiziqlari 5° dan katta xatolar sohasida o‘qituvchi egri chizig‘idan sezilarli darajada yuqorida joylashadi, ya’ni ular tasvirlarning katta qismini aniqroq baholaydi. KD o‘quvchi egri chizig‘i esa deyarli butun oraliqda oddiy o‘quvchi egri chizig‘idan yuqorida yoki u bilan ustma-ust joylashadi.

O‘quvchi modellar o‘qituvchi modeldan taxminan 39–44 marta tez ishlaydi. KD va oddiy o‘quvchilarning arxitekturasi bir xil, shu sababli ular orasidagi kichik farq o‘lchash jarayonidagi tebranishlar bilan izohlanadi. TorchScript formatiga eksport qilingan model fayli 0,60 MiB ni tashkil etadi, qayta yuklangan model bilan asl model bashoratlari orasidagi maksimal farq esa 0 ga teng.

Taqqoslash uchun, MPIIGaze to‘plamining to‘liq hajmida «bir ishtirokchini chiqarib qoldirish» (leave-one-person-out) protokoli bo‘yicha [3] ishda MnistNet modeli uchun 6,3°, GazeNet modeli uchun 5,5° o‘rtacha xato qayd etilgan. Ushbu ishda tasvirlar soni (1500 ta) va baholash protokoli (2 ta test ishtirokchisi) farq qilgani sababli olingan 7,05° natijani ular bilan bevosita taqqoslab bo‘lmaydi.

E’tiborli jihati shundaki, o‘qituvchi model o‘rtacha aniqlik bo‘yicha o‘quvchilardan ortda qolgan bo‘lsa-da, uning bashoratlari o‘quvchi model aniqligini oshirdi. Bizning fikrimizcha, bu holat o‘qituvchi bashoratlarining qo‘shimcha regulyarizatsiya vazifasini bajarishi bilan izohlanadi: o‘qituvchi eng aniq baholagan pitch burchagi bo‘yicha KD o‘quvchining xatosi 4,23° dan 3,75° gacha kamaydi. Demak, bilim distillatsiyasidan foyda olish uchun o‘qituvchi model barcha ko‘rsatkichlar bo‘yicha ustun bo‘lishi shart emas.

Olingan natijalar asosida quyidagi amaliy tavsiyalar ishlab chiqildi. Birinchidan, qurilmada ishlovchi ko‘z kuzatuvi tizimlarini yaratishda yirik vizual modellarni bevosita joylashtirish o‘rniga ulardan o‘qituvchi model sifatida foydalanish va bilim distillatsiyasi orqali kichik model olish maqsadga muvofiq. Ikkinchidan, qarash yo‘nalishini baholovchi modellarni sinashda ma’lumotlarni ishtirokchilar bo‘yicha ajratish majburiy talab sifatida qo‘llanilishi lozim, aks holda modelning yangi foydalanuvchilardagi aniqligi yuqori baholanib yuborilishi mumkin. Uchinchidan, distillatsiya koeffitsiyenti α validatsiya to‘plami bo‘yicha tanlanishi kerak, kichik hajmli to‘plamlar uchun $\alpha = 0{,}3$ qiymatini ablatsiya tajribasi bilan tekshirilguniga qadar boshlang‘ich qiymat sifatida qo‘llash mumkin. To‘rtinchidan, olingan 0,60 MiB hajmli model haydovchi hushyorligini nazorat qiluvchi mobil ilovalarga [11] qo‘shimcha modul sifatida integratsiya qilinishi mumkin: qarash yo‘nalishi haydovchining yo‘ldan chalg‘ishini aniqlashda ko‘z yumilishi va esnash belgilarini to‘ldiruvchi muhim belgi bo‘lib xizmat qiladi.

Shu bilan birga, test qismi ikki ishtirokchidan iborat va taqqoslash bitta tasodifiy urug‘ bilan bajarilganini hisobga olish lozim. Bundan tashqari, leave-one-person-out protokoli

DOI: 10.65322/j.srd.05192026.128

qo‘llanmagan, α koeffitsiyenti ablatsiya bilan asoslanmagan, tezlik esa faqat CPU da o‘lchangan va mobil qurilmada tekshirilmagan. Shu sababli kengroq xulosa uchun tajribani ko‘proq ishtirokchilar, bir nechta tasodifiy urug‘ va real kamera oqimida takrorlash talab etiladi.

Xulosa. Ushbu tadqiqotda yirik vizual modeldan bilim distillatsiyasi orqali ko‘z qarash yo‘nalishini baholovchi yengil model ishlab chiqildi va MPIIGaze to‘plamining 1500 ta tasvirida ishtirokchilar bo‘yicha ajratish sharoitida sinovdan o‘tkazildi. Bilim distillatsiyasi o‘quvchi modelning o‘rtacha burchak xatosini 7,38° dan 7,05° gacha kamaytirdi va xatosi 5° dan kichik tasvirlar ulushini 27,5 % dan 34,5 % gacha oshirdi, bunda model hajmi (146 786 parametr) va CPU dagi kechikish (~0,26 ms) o‘zgarmadi.

Mazkur tadqiqotning ilmiy ahamiyati shundaki, u o‘tkazilgan tajriba sharoitida bilim distillatsiyasi o‘qituvchi model barcha ko‘rsatkichlar bo‘yicha ustun bo‘lmagan holatda ham kichik modelning aniqligini uning hajmi va tezligini o‘zgartirmagan holda oshirishi mumkinligini ko‘rsatdi. Amaliy ahamiyati esa olingan modelning smartfon va avtomobil bortidagi qurilmalarda real vaqt rejimida ishlovchi ko‘z kuzatuvi tizimlari uchun istiqbolli ekanidadir.

Kelgusida tajribani ko‘proq ishtirokchilar va bir nechta tasodifiy urug‘ bilan leave-one-person-out protokoli bo‘yicha takrorlash, α koeffitsiyenti bo‘yicha ablatsiya o‘tkazish, model kirishini real kamera oqimidan ko‘zni aniqlash moduli bilan bog‘lash hamda mobil qurilmada tezlikni o‘lchash rejalashtirilmoqda.

Foydalanilgan adabiyotlar ro‘yxati

1. Jiang C. va boshq. Rapidly deploying on-device eye tracking by distilling visual foundation models // European Conference on Computer Vision (ECCV). – 2026. – arXiv:2604.02509.

2. Zhang X., Sugano Y., Fritz M., Bulling A. Appearance-based gaze estimation in the wild // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2015. – P. 4511–4520.

3. Zhang X., Sugano Y., Fritz M., Bulling A. MPIIGaze: Real-world dataset and deep appearance-based gaze estimation // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2019. – Vol. 41, No. 1. – P. 162–175.

4. Krafka K., Khosla A., Kellnhofer P., Kannan H., Bhandarkar S., Matusik W., Torralba A. Eye tracking for everyone // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – P. 2176–2184.

5. Hinton G., Vinyals O., Dean J. Distilling the knowledge in a neural network // arXiv:1503.02531. – 2015.

6. Dosovitskiy A., Beyer L., Kolesnikov A. va boshq. An image is worth 16x16 words: Transformers for image recognition at scale // International Conference on Learning Representations (ICLR). – 2021.

7. Oquab M., Darcet T., Moutakanni T. va boshq. DINOv2: Learning robust visual features without supervision // Transactions on Machine Learning Research. – 2024. – arXiv:2304.07193.

8. Bulling A. MPIIGaze [ma’lumotlar to‘plami] // DaRUS, V1. – 2022. – DOI: 10.18419/DARUS-3230.

9. Hoerl A.E., Kennard R.W. Ridge regression: Biased estimation for nonorthogonal problems // Technometrics. – 1970. – Vol. 12, No. 1. – P. 55–67.

10. Paszke A., Gross S., Massa F. va boshq. PyTorch: An imperative style, high-performance deep learning library // Advances in Neural Information Processing Systems 32 (NeurIPS 2019). – 2019. – P. 8024–8035.

11. Xolmurodov S.L., Safarov F.E. Real vaqt rejimida haydovchi uyquchanligini aniqlash muammosi va mavjud usullar tahlili // Sun’iy intellekt texnologiyalarining fan va ta’limda joriy etish tajribasi: muammo va istiqbollari: Respublika ilmiy-amaliy anjuman materiallari. – Navoiy, 2026. – B. 15–20.