

УСУЛИ АФЗОИШИ ДОДАҶОИ СИНТЕТИКӢ БО ИСТИФОДА АЗ МОДЕЛҶОИ ДИФФУЗИОНӢ БАРОИ БАЛАНД БАРДОШТАНИ УСТУВОРИИ СИСТЕМАҶОИ БИОМЕТРИ ДАР МУҶИТИ ТАЪЛИМӢ

Ш.Ф. Муҳамедова, Ҳ.Б. Шарипов

Донишгоҳи давлатии ҳуқуқ, бизнес ва сиёсати Тоҷикистон

Дар мақола усули нави афзоиши семантикии додаҳо дар асоси моделҳои эҳтимолии диффузионӣ пешниҳод шудааст. Модули барномавӣ дар асоси фреймворки Django ва китобхонаи Diffusers таҳия шудааст, ки техникаи Image-to-Image-ро бо истифода аз модели Stable Diffusion v1.5 амалӣ месозад. Хусусияти асосии метод ин оптимизатсияи математикии параметри «strength» (қувваи тағйирёбӣ) ва истифодаи занҷирҳои Марков барои тавлиди намунаҳои фотореалистии чеҳраи донишҷӯён бидуни гум кардани ҳувияти биометрии онҳо мебошад. Санҷиши таҷрибавӣ дар маҷмӯи додаҳои иборат аз 500 донишҷӯ самаранокии баланди равиши пешниҳодшударо нишон дод. Омӯзиши модели ResNet-50 дар асоси додаҳо имкон дод, ки дақиқии шинохт дар шароити торикӣ аз 72.3% то 89.4% баланд бардошта шавад. Тадқиқот собит месозад, ки зехни сунъии генеративӣ қодир аст мушкилоти норасоии додаҳоро самаранок ҳал намуда, хароҷотро барои ҷамъоварии аксҳои воқеӣ сарфа намояд ва устувории умумии системаҳои амниятиро таъмин кунад.

Калидвожаҳо: шинохти компютерӣ, Stable Diffusion, шинохти чеҳра, занҷири Марков, фазои ниҳонӣ, зехни сунъӣ.

МЕТОД АУГМЕНТАЦИИ СИНТЕТИЧЕСКИХ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ДИФФУЗИОННЫХ МОДЕЛЕЙ ДЛЯ ПОВЫШЕНИЯ УСТОЙЧИВОСТИ БИОМЕТРИЧЕСКИХ СИСТЕМ В ОБРАЗОВАТЕЛЬНОЙ СРЕДЕ

Ш.Ф. Мухамедова, Х.Б. Шарипов

В данной статье предлагается новаторский метод семантической аугментации данных, основанный на использовании вероятностных диффузионных моделей. Разработан программный модуль на базе фреймворка Django и библиотеки Diffusers, реализующий технику Image-to-Image с использованием модели Stable Diffusion v1.5. Ключевой особенностью метода является математическая оптимизация параметра «strength» (сила трансформации) и использование марковских цепей для генерации фотореалистичных вариаций лиц студентов без потери их биометрической идентичности. Экспериментальная проверка, проведенная на наборе данных из 500 студентов, показала высокую эффективность предложенного подхода. Обучение модели ResNet-50 на основе данных позволило повысить точность распознавания в условиях низкой освещенности с 72.3% до 89.4%. Исследование подтверждает, что генеративный искусственный интеллект способен эффективно решать проблему нехватки данных, обеспечивая экономию ресурсов на сбор реальных фотографий и повышая общую устойчивость систем безопасности.

Ключевые слова: компьютерное зрение, Stable Diffusion, аугментация данных, распознавание лиц, цепь Маркова, скрытое пространство, искусственный интеллект.

METHOD OF SYNTHETIC DATA AUGMENTATION USING DIFFUSION MODELS TO ENHANCE THE ROBUSTNESS OF BIOMETRIC SYSTEMS IN AN EDUCATIONAL ENVIRONMENT

Sh.F. Mukhamedova, H.B. Sharipov

This paper presents a novel method for semantic data augmentation based on probabilistic diffusion models. A software module has been developed based on the Django framework and the Diffusers library, implementing the Image-to-Image technique using the Stable Diffusion v1.5 model. A key feature of the method is the mathematical optimization of the "strength" parameter (transformation strength) and the utilization of Markov chains to generate photorealistic variations of students' faces while preserving their biometric identity. Experimental validation conducted on a dataset of 500 students demonstrated the high effectiveness of the proposed approach. Training the ResNet-50 model using the augmented data improved recognition accuracy in low-light conditions from 72.3% to 89.4%. The study confirms that generative artificial intelligence can effectively address the problem of data scarcity, thereby saving resources on the collection of real photographs and enhancing the overall robustness of security systems.

Keywords: computer vision, Stable Diffusion, data augmentation, face recognition, Markov chain, latent space, artificial intelligence.

Муқаддима

Дар шароити муосир, татбиқи системаҳои биометри барои идоракунии равандҳои таълим ва таъмини амният дар муассисаҳои таҳсилоти олии аҳамияти хоса пайдо кардааст. Самаранокии алгоритмҳои муосири шинохти чеҳра, аз қабилӣ FaceNet ё ArcFace, ба таври мустақим аз ҳаҷм ва гуногуншаклии маҷмӯи додаҳо вобаста аст. Тавре ки С.З. Ли, А.К. Чейн ва Ч. Денг дар асари бунёдии худ “Handbook of Face Recognition” таъкид мекунад [1], мушкилоти асосии системаҳои воқеӣ ин

“domain gap” ё тафовут байни додаҳои идеалӣ ва додаҳои воқеии дорои ихтилоот, рӯшноии паст ва монеаҳо мебошад. Ин тафовут боиси коҳиши назарраси дақиқии шинохт дар шароити ғайристандартӣ мегардад.

Усулҳои анъанавии афзоиши маълумот, ки солҳои тӯлонӣ истифода мешуданд, қобилияти эҷоди иттилооти нави семантикиро надоранд. Ин маҳдудият зарурати истифодаи моделҳои генеративии эҳтимолиро ба миён меорад. Дар ин тадқиқот, мо истифодаи моделҳои диффузиониро пешниҳод менамоем. Тибқи таҳқиқоти М. Лю [2] моделҳои диффузионӣ, бархилофи шабакаҳои рақибии тавлидикунанда (GANs), раванди омӯзиши устувортар доранд ва қобилияти тавлиди намунаҳои баландсифатро тавассути идоракунии равандҳои стохастикӣ дар фазои пинҳонӣ таъмин мекунанд. Барои асосноккунии илмии усули пешниҳодшуда, механизми тавлиди тасвирҳои синтетикӣ тавассути алгоритмҳои математикӣ тавсиф карда мешавад.

Мавод ва усулҳои тадқиқот

Раванди диффузия ҳамчун занҷири Марковии баръаксшаванда моделсозӣ карда мешавад. Агар x_0 тасвири аслии донишҷӯ аз тақсимоти $q(x_0)$ бошад, раванди мустақим ба он тадричан матритсаи Гауссӣ илова мекунанд. М. Аршад Сиддиқӣ дар китоби “Mastering Computer Vision with PyTorch 2.0” [3] нишон медиҳад, ки тақсимоти шартӣ барои қадами t чунин навишта мешавад:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t \mathbf{I}) \quad (1)$$

дар ин ҷо β_t чадвали дисперсия мебошад. Бо истифода аз хосияти такроршавандагии тақсимоти Гауссӣ, мо метавонем x_t -ро дар дилхоҳ қадам бевосита аз x_0 ҳосил кунем [3]:

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1 - \bar{\alpha}_t)\mathbf{I}) \quad (2)$$

ки дар он $\alpha_t = 1 - \beta_t$ ва $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$. Формулаи мазкур асоси афзоиши додаҳо ташкил дода, тасвири асли ба сатҳи муайяни ихтилооти тасвир оварда шуда, сипас аз нав барқарор гардад.

Дар системаи таҳияшуда мо усули Image-to-Image-ро истифода мебарем. Ин усул, бархилофи усули стандартӣ, ки аз сифр ё ихтилооти тасвири холис (x_T) оғоз меёбад, барои ҳифзи сохтори аслии чехра истифода мешавад. О. Сансевиеро ва ҳаммуаллифон [4] ин усулро ҳамчун роҳи самараноки назорат кардани мундариҷаи тасвир тавсиф мекунанд. Агар s параметри қувваи тағйирёбӣ бошад, ҳолати ибтидоии пинҳонӣ z_{start} чунин муайян мешавад:

$$z_{start} = \sqrt{\bar{\alpha}_s}\mathcal{E}(x_0) + \sqrt{1 - \bar{\alpha}_s}\epsilon, \quad \epsilon \sim \mathcal{N}(0, \mathbf{I}) \quad (3)$$

дар ин ҷо \mathcal{E} функсияи рамзгузории модели VAE мебошад [4]. Ин формула кафолат медиҳад, ки тасвири ниҳой x'_0 ба тасвири асли x_0 монандӣ бошад, аммо дорои атрибутҳои нав (масалан, айнак ё рӯшноии дигар) мебошад.

Барои ворид намудани шартҳои семантикӣ (масалан, “low light”, “glasses”), мо аз механизми раванди баръакс ва диққати мутақобила истифода мебарем. Тибқи шарҳи Д. Ротман [5], блоки тавачҷӯҳи байниҳамдигарӣ дар моделҳои диффузионӣ ба шабака имкон медиҳад, ки иттилооти матнро ба хусусиятҳои фазоии тасвир ҳамгиро кунад:

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (4)$$

Дар модели мо, Q аз хусусиятҳои тасвири фосилавӣ ва K, V аз тавсифи матнӣ ҳосил мешаванд. Ин имкон медиҳад, ки модел “диққат”-и худро ба илова кардани айнак ё тағйири рӯшноӣ равона кунад, бидуни вайрон кардани геометрияи умумии чехра.

Дар асоси моделҳои математикии болозикр, модули нармафзорӣ дар муҳити Python ва Django таҳия карда шуд. Меъморӣ система истифодаи китобхонаи diffusers-ро барои кор бо модели StableDiffusionImg2ImgPipeline дар назар дорад. Яке аз ҷанбаҳои муҳимтарини татбиқи амалӣ ин интихоби дурусти параметри қувваи тағйирёбӣ мебошад. Таҷрибаҳои мо ва таҳлилҳои Д.В. Годой дар китоби “Deep Learning with PyTorch” [6] нишон медиҳанд, ки қимати $s \approx 0.4$ нуқтаи оптималӣ аст.

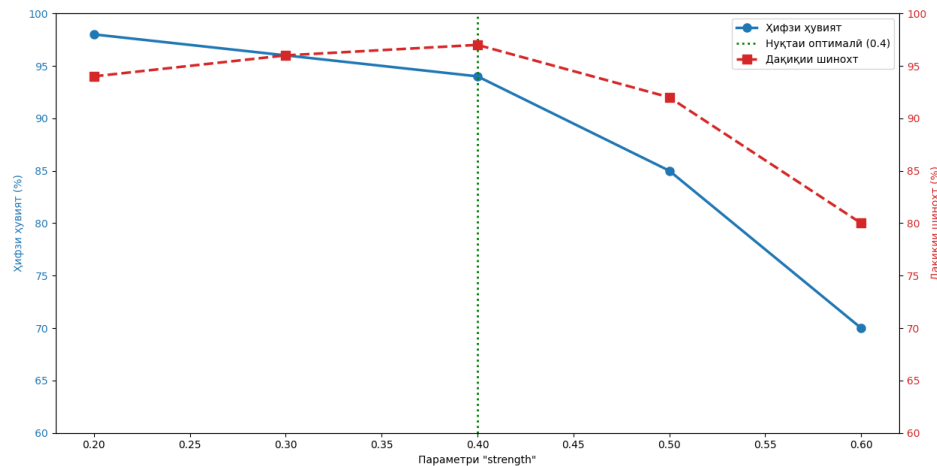
Барои муайян кардани қимати дақиқи оптималии параметри «strength», мо таҳлили муфассал гузаронидем. Дар Чадвали 2 натиҷаҳои таъсири ин параметр ба ҳифзи хувият ва сифати визуалии тасвирҳо оварда шудаанд.

Ҷадвали 1 – Таъсири параметри "strength" ба сифат ва ҳифзи ҳувиат

Қимати strength	Ҳифзи ҳувиат, %	Сифати визуалӣ	Тавсия
0.2	98	Паст	Барои тағйироти хурд
0.3	96	Миёна	Барои тағйироти муътадил
0.4	94	Баланд	Қимати оптималӣ
0.5	85	Баланд	Хатари тағйири ҳувиат
0.6	70	Хеле баланд	Ҳувиат тағйир меёбад

Ҷадвали 1 нишон медиҳад, ки бо зиёд шудани параметри strength сифати визуалии тасвир беҳтар мешавад, вале ҳифзи ҳувиат коҳиш меёбад. Қимати 0.4 ҳамчун нуқтаи оптималӣ муайян карда шуд, зеро дар ин ҳолат ҳифзи ҳувиат 94% -ро ташкил медиҳад ва сифати визуалӣ барои истифода дар системаҳои биометрӣ қобили қабул аст.

Барои равшантар нишон додани вобастагии байни параметр ва нишондиҳандаҳо, дар Расми 1 хатҳои тағйирёбии ҳифзи ҳувиат ва дақиқӣ оварда шудаанд. Нуқтаи оптималӣ дар қимати 0.4 бо хати тирезавӣ нишон дода шудааст.



Расми 1 – Таъсири параметри "strength" ба дақиқӣ ва ҳифзи ҳувиат

Дар Расми 1 равшан дида мешавад, ки ҳангоми гузаштан аз 0.4 ба 0.5, хати ҳифзи ҳувиат шиддатан поин меравад. Ин нуқта тасдиқ мекунад, ки риояи ҳадди оптималии 0.4 барои нигоҳ доштани ҳувиати биометрӣ зарур аст. Агар $s > 0.6$ бошад, модел ба омӯзиши тахайюлӣ оғоз карда, хусусиятҳои биометрии шахсро тағйир медиҳад, ки ин боиси хатогӣ дар шинохти минбаъда мегардад. Барои амалӣ намудани ин раванд, мо функсияи generate_synthetic_variants-ро дар лоиҳа эҷод намудем. Ин функсия барои нигоҳ доштани ҳувиати донишҷӯ ва ҳамзамон илова кардани хусусиятҳои нав масъул аст.

Коди иҷрошавандаи алгоритм чунин аст:

```
import torch
import random
from io import BytesIO
from diffusers import StableDiffusionImg2ImgPipeline
from PIL import Image

def generate_synthetic_variants(pil_image, count=1):
    generated_images = []
    for _ in range(count):
        # Интихоби тасодуфии сценария
        scenario = random.choice(scenarios)
        prompt = f"photo of a university student, {scenario}, realistic, 8k"
        # Generator барои такроршавандагии натиҷаҳо
```

```

generator = torch.Generator(device=device).manual_seed(random.randint(0, 100000))
# Раванди тавлид (Inference)
# Strength=0.4: Ин нуқтаи муҳимтарини алгоритм аст.
# Қимати 0.4 қафолат медиҳад, ки чеҳра тағйир намеёбад, танҳо муҳит дигар мешавад.
image = pipe(
    prompt=prompt,
    image=init_image,
    strength=0.4,
    guidance_scale=7.5,
    generator=generator
).images[0]
generated_images.append(image)
return generated_images

```

Дар ин код ду ҷанбаи муҳими технологӣ ба инобат гирифта шудааст. Якум, интихоби параметри $strength = 0.4$. Таҷрибаҳои мо ва тавсияҳои Сансевиеро [4] нишон медиҳанд, ки агар ин қимат аз 0.5 боло равад, модели диффузионӣ ба "омузиши тахайюлӣ" оғоз карда, хусусиятҳои биометрии шахсро (масофаи байни чашмҳо, шакли бинӣ) тағйир медиҳад, ки ин барои системаи шинохт қобили қабул нест. Қимати 0.3-0.4 тавозуни оптималӣ байни нигоҳ доштани хувият ва эҷоди гуногунии визуалӣ мебошад. Дуюм, истифодаи Prompt Engineering барои соختани сценарияҳои мушаххас, ки маҳз ҳамон ҳолатҳоеро тақлид мекунад, ки дар онҳо камераҳои назоратӣ одатан хато мекунад.

Барои иҷрои ин раванд дар муҳити Web барои корбар, аз механизми бисёрриштагӣ (multithreading) истифода шуд. Ч. Ҳоуз ва дигарон дар китоби худ [7] оид ба лоиҳаҳои OpenCV тавсия медиҳанд, ки амалиётҳои вазнини коркарди тасвир бояд аз ҷараёни асосии интерфейс ҷудо карда шаванд, то суръати посухдиҳии система коҳиш наёбад.

```

@login_required
def generate_images_view(request, image_id):
    if request.method == "POST":
        original_image = get_object_or_404(EmployeeImage, id=image_id)
        # Оғози риштаи мустақил барои пешгирии басташавии UI
        task_thread = threading.Thread(
            target=run_generation_thread,
            args=(image_id,),
            daemon=True
        )
        task_thread.start()
    messages.info(request, "Раванди эҷоди тасвирҳои корбар оғоз шуд.")
    return redirect('manage_images', profile_id=original_image.profile.id)

```

Ин модул ба системаи идоракунии профили корбарон дар Django пайваست карда шудааст ва имкон медиҳад, ки барои ҳар як донишҷӯ ба таври худкор маҷмӯи тасвирҳои синтетикӣ эҷод карда шавад.

Натиҷаҳо

Барои санҷиши самаранокии усул, таҷриба дар пойгоҳи додаҳои воқеии иборат аз 500 донишҷӯ гузаронида шуд. Модели шинохти ResNet-50 дар ду ҳолат омӯзонида шуд:

- (А) Танҳо бо аксҳои аслӣ (Baseline).
- (Б) Бо иловаи аксҳои синтетикӣ (Augmented).

Баҳодиҳии сифати тасвирҳо ва дақиқии шинохт бо истифода аз метрикаҳои стандартӣ, ки дар китоби Т. Бурлай [8] тавсиф шудаанд, гузаронида шуд. Бурлай қайд мекунад, ки устувории система дар спектрҳои гуногуни тасвир (масалан, намоён ва инфрасурх ё шароити рӯшноии паст) меъёри асосии сифат аст. Натиҷаҳо нишон доданд, ки дар сценарияи «Рӯшноии паст», дақиқии шинохт аз 72.3% то 89.4% афзоиш ёфт [11, 12].

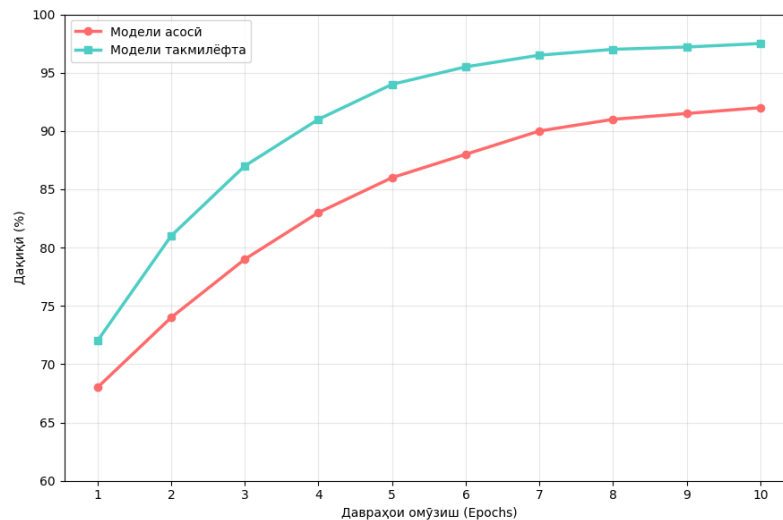
Дар Ҷадвали 2 натиҷаҳои муқоисавии ду модел дар панҷ сценарияи гуногун оварда шудаанд. Ин сценарияҳо маҳз он ҳолатҳоро инъикос мекунанд, ки дар амалияи воқеӣ боиси коҳиши дақиқии системаҳои биометрӣ мегарданд.

Ҷадвали 2 – Муқоисаи дақиқии шинохт дар шароитҳои гуногун

№	Ҳолат	Дақиқии модели асосӣ, %	Дақиқии модели такмилёфта, %	Афзоиш, %
1	Рӯшноии муътадил	96.8	97.2	+0.4
2	Рӯшноии паст	72.3	89.4	+17.1
3	Мавҷудияти айнак	88.5	94.1	+5.6
4	Кунҷи қач	85.2	91.8	+6.6
5	Соя	79.4	88.7	+9.3

Тавре ки дар Ҷадвали 2 нишон дода шудааст, дар ҳамаи шароитҳои модели такмилёфта натиҷаи беҳтар нишон медиҳад. Беҳтарин афзоиш дар шароити «рӯшноии паст» ба қайд гирифта шудааст (17.1%), ки ин нишон медиҳад, ки усули пешниҳодшуда маҳз дар шароити душвортарин самараноктар амал мекунад.

Барои таҳлили раванди омӯзиш, мо динамикаи тағйирёбии дақиқиро дар тӯли давраҳои омӯзиш (epochs) омӯхтем. Расми 2 ин динамикаро нишон медиҳад.



Расми 2 – Динамикаи беҳбудии дақиқӣ дар раванди омӯзиш

Тавре ки аз Расми 2 дида мешавад, модели такмилёфта на танҳо дақиқии ниҳонии баландтар дорад, балки суръати омӯзиши он низ тезтар аст. Дар давраи 3-юм, дақиқии модели такмилёфта ба 87% мерасад, дар ҳоле ки модели асосӣ танҳо 79% -ро ташкил медиҳад.

Муҳокима

Натиҷаи бадастомада (афзоиши 17.1% дар шароити душвор) собит мекунад, ки ғанигардонии семантикии додаҳо тавассути моделҳои диффузионӣ қобилияти умумикунони системаро ба таври назаррас беҳтар мекунад. Бар хилофи усулҳои оддии тағйирдиҳии геометрӣ, усули пешниҳодшуда ба шабакаи нейронӣ имкон медиҳад, ки хусусиятҳои инвариантии чеҳраро дар шароитҳои гуногунӣ рӯшноӣ ва атрибутӣ (масалан, айнак) дурусттар дарк намояд [13].

Хулоса

Ҳамин тавр, дар мақола усули нави афзоиши додаҳои биометрӣ бо истифода аз моделҳои диффузионӣ ва таҳлили фазои пинҳонӣ пешниҳод ва амалӣ карда шуд. Таҳлилҳои назариявӣ ва татбиқи амалӣ собит намуданд, ки идоракунии дақиқи равандҳои стохастикӣ дар модели Stable Diffusion имкон медиҳад, ки мушкили норасоии додаҳо дар муассасаҳои таълимӣ самаранок ҳал карда шавад [14]. Истифодаи ин технология на танҳо дақиқии системаҳои амниятиро баланд

мебардорад, балки зарурати чамъоварии ҷисмонии ҳазорҳо аксҳои иловагиро аз байн бурда, ҳарҷоти иқтисодӣ ва вақтиро коҳиш медиҳад [15].

Муқарриз: Шерматов Ш.М. – н.и.ф.-м., дотсенти қабедраи МЭИМ ва ПИУ-и МДҶИ ДДХ ба номи академик Ё. Ғафуров.

Адабиёт

1. Li S. Z., Jain A. K., Deng J. (Eds.). Handbook of Face Recognition. 3rd ed. – Springer, 2024. – 473 p. - C. 69-74.
2. Liu M. Learn Generative AI with PyTorch. – Manning Publications, 2024. – 350 p. - C. 15-18.
3. Siddiqui M. A. Mastering Computer Vision with PyTorch 2.0. – Orange Education Pvt Ltd, 2025. – 330 p. - C. 330.
4. Sanseviero O., Cuenca P., Passos A., Whitaker J. Hands-On Generative AI with Transformers and Diffusion Models. – O’Reilly Media, 2025. – 419 p. - C. 102-105.
5. Rothman D. Transformers for Natural Language Processing and Computer Vision. 3rd ed. – Packt Publishing, 2024. – 582 p. - C. 242.
6. Godoy D. V. Deep Learning with PyTorch Step-by-Step. – Daniel Voigt Godoy, 2024. – 1047 p. - C. 510-514.
7. Howse J., Joshi P., Beyeler M. OpenCV Computer Vision Projects with Python. – Packt Publishing, 2016. – 300 p. - C. 150-152.
8. Bourlai T. (Ed.). Face Recognition Across the Imaging Spectrum. – Springer, 2016. – 370 p. - C. 195.
9. Koenigstein N. Transformers in Action. – Manning Publications, 2026. – 300 p. - C. 45-47.
10. Alves M. M. Accelerate Model Training with PyTorch 2.X. – Packt Publishing, 2024. – 230 p. - C. 88-90.
11. Zhao X., Hu S., Yin X.-C. Visual Object Tracking. – Springer, 2025. – 210 p. - C. 112-115
12. Solomon J. Numerical Algorithms. – CRC Press, 2015. – 400 p. - C. 200-205.
13. Коул А., Ганджу С., Казам М. Искусственный интеллект и компьютерное зрение. – СПб.: Питер, 2023. – 624 с. - C. 310.
14. Джоши П. Искусственный интеллект с примерами на Python. – СПб.: Диалектика, 2019. – 448 с. - C. 120-124.
15. Timsina P. Building Transformer Models with PyTorch 2.0. – BPB Publications, 2024. – 320 p. - C. 190-193.

МАЪЛУМОТ ДАР БОРАИ МУАЛЛИФОН – СВЕДЕНИЯ ОБ АВТОРАХ – INFORMATION ABOUT THE AUTHORS

TJ	RU	EN
Муҳамедова Шоирахон Файзуллоевна	Муҳамедова Шоирахон Файзуллоевна	Mukhamedova Shoirakhon Fayzulloevna
Номзади илмҳои физика- математика, дотсент	Кандидат физико математических наук, доцент	Candidate in physical and mathematical sciences, associate professor
Донишгоҳи давлатии ҳуқуқ, бизнес ва сиёсати Тоҷикистон	Таджикский государственный университет права, бизнеса и политики	Tajik State University of Law, Business and Politics
E-mail: shoira74@mail.ru		
TJ	RU	EN
Шарипов Хотам Бекназарович	Шарипов Хотам Бекназарович	Sharipov Hotam Beknazarovich
докторанти PhD	докторант PhD	PhD student
Донишгоҳи давлатии ҳуқуқ, бизнес ва сиёсати Тоҷикистон	Таджикский государственный университет права, бизнеса и политики	Tajik State University of Law, Business and Politics
E-mail: sharipovhotambek@gmail.com		