O‘ZBEKISTON MILLIY UNIVERSITETI XABARLARI, 2026, [9]
ISSN 2181-7324(Print)

IJTIMOIY FANLAR
http://journals.nuu.uz
ISSN 3093-8996(Onlayn)

Oybek RAJABOV,
Katta o‘qituvchi, Chirchiq davlat pedagogika universiteti, Chirchiq, O‘zbekiston
E-mail: o.rajabov@cspu.uz, ORCID:0000-0002-2966-3807

PhD M.Berdimurodov taqrizi asosida

FORMATION OF PRELIMINARY DATASET FOR UZBEK SIGN LANGUAGE AND DEVELOPMENT OF RECOGNITION MODEL BASED ON ARTIFICIAL INTELLIGENCE

Annotation

This research addresses the formation of a preliminary dataset for Uzbek Sign Language and the development of an AI-based recognition model to ensure digital inclusion of persons with hearing and speech impairments in the Republic of Uzbekistan. The relevance stems from the President’s "Digital Uzbekistan – 2030" strategy and state programs supporting persons with disabilities. The methodology employed hand landmark extraction via MediaPipe, temporal sequence processing using LSTM architecture, and comparative framework analysis. As a result, 15,000 video samples for 50 gestures were collected from 30 participants, with the model achieving 80.7% accuracy on the test set. The scientific novelty lies in creating the first open dataset prototype accounting for Uzbek sign language specifics, while practical significance is demonstrated in enabling future real-time sign language interpreter systems.

Key words: Sign language, dataset creation, artificial intelligence, LSTM, gesture classification, Uzbek sign language, database, MediaPipe, time-series classification, deep learning.

ФОРМИРОВАНИЕ ПРЕДВАРИТЕЛЬНОЙ БАЗЫ ДАННЫХ ДЛЯ УЗБЕКСКОГО ЯЗЫКА ЖЕСТОВ И РАЗРАБОТКА МОДЕЛИ РАСПОЗНАВАНИЯ НА ОСНОВЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Аннотация

В данном исследовании рассматриваются вопросы формирования предварительной базы данных для узбекского языка жестов и разработки модели распознавания на основе искусственного интеллекта с целью обеспечения цифровой инклюзии лиц с нарушениями слуха и речи в Республике Узбекистан. Актуальность обусловлена стратегией Президента «Цифровой Узбекистан – 2030» и государственными программами поддержки лиц с ограниченными возможностями. В качестве методологии использовались извлечение контрольных точек руки через MediaPipe, обработка временных рядов с помощью архитектуры LSTM и сравнительный анализ фреймворков. В результате собрано 15 000 видеозаписей для 50 жестов от 30 участников, модель показала точность 80,7% на тестовой выборке. Научная новизна заключается в создании первого открытого прототипа датасета с учётом особенностей узбекского языка жестов, а практическая значимость — в возможности создания системы сурдоперевода в реальном времени.

Ключевые слова: Язык жестов, создание датасета, искусственный интеллект, LSTM, классификация жестов, узбекский язык жестов, база данных, MediaPipe, классификация временных рядов, глубокое обучение.

O‘ZBEK IMO-ISHORA TILI UCHUN DASTLABKI MA’LUMOTLAR BAZASINI SHAKLLANTIRISH VA SUN’IY INTELLEKT ASOSIDA TANIB OLISH MODELINI ISHLAB CHIQISH

Annotatsiya

Ushbu tadqiqotda O‘zbekiston Respublikasida eshitish va nutqida nuqsoni bo‘lgan shaxslarning raqamli inkluziyasini ta’minlash maqsadida o‘zbek imo-ishora tili uchun dastlabki ma’lumotlar bazasi (dataset) shakllantirish va sun’iy intellekt asosida tanib olish modelini ishlab chiqish masalalari yoritilgan. Tadqiqot dolzarbligi Prezidentimiz Sh.M. Mirziyoyevning “Raqamli O‘zbekiston – 2030” strategiyasi va nogironligi bo‘lgan shaxslarni qo‘llab-quvvatlash bo‘yicha davlat dasturlaridan kelib chiqadi. Metodologiya sifatida qo‘l landmarklarini MediaPipe orqali ekstraksiya qilish, vaqt seriyali ma’lumotlarni LSTM arxitekturasi bilan qayta ishlash va frameworklar taqqoslash usullari qo‘llanildi. Natijada 30 nafar ishtirokchidan 50 ta imo-ishora uchun 15,000 video yozuv yig’ildi, model test bosqichida 80.7% aniqlik ko‘rsatdi. Tadqiqotning ilmiy yangiligi o‘zbek imo-ishora tili xususiyatlarini hisobga olgan birinchi ochiq dataset prototipini yaratishda, amaliy ahamiyati esa kelgusida real vaqt rejimida ishlovchi surdo-tarjimon tizimini yaratish imkoniyatida namoyon bo‘ladi.

Kalit so‘zlar: imo-ishora tili, ma’lumotlar to‘plamini yaratish, sun’iy intellekt, LSTM modeli, imo-ishoralarni tasniflash, o‘zbek imo-ishora tili, ma’lumotlar bazasi, MediaPipe texnologiyasi, vaqt qatorlari tasnifi, chuqur o‘rganish.

Kirish. O‘zbekiston Respublikasida raqamli transformatsiya jarayonlarini chuqurlashtirish va aholining barcha qatlamlari uchun teng imkoniyatlar yaratish Prezidentimiz Shavkat Mirziyoyev tomonidan belgilangan ustuvor vazifalardan biridir. Xususan, “Raqamli O‘zbekiston – 2030” strategiyasida axborot-kommunikatsiya texnologiyalari orqali ijtimoiy adolatni ta’minlash, nogironligi bo‘lgan shaxslarning ta’lim, mehnat va ijtimoiy hayotda faol ishtirokini rag‘batlantirish alohida ta’kidlangan [1]. Eshitish va nutqida nuqsoni bo‘lgan fuqarolar uchun imo-ishora tili ularning kundalik muloqot, ta’lim olish va kasb tanlashdagi asosiy vositasi hisoblanadi.

Ammo bugungi kunda o‘zbek imo-ishora tili rasmiy ravishda raqamlashtirilmagan, standartlashtirilmagan va ochiq ma’lumotlar bazalariga ega emas. Bu holat sun’iy intellekt (AI) asosidagi yordamchi texnologiyalarni mahalliylashtirishda jiddiy ilmiy va texnik to‘siq bo‘lib qolmoqda. Dunyo tajribasi shuni ko‘rsatadiki, imo-ishora tilini avtomatik tanib olish

tizimlarining samaradorligi to‘g‘ridan-to‘g‘ri keng ko‘lamli, sifatli va lingvistik jihatdan asoslangan ma‘lumotlar bazasiga bog‘liq [2].

Shu bois, ushbu tadqiqotning asosiy maqsadi – O‘zbek imo-ishora tili uchun birinchi dastlabki, sifatli va metodologik jihatdan asoslangan ma’lumotlar bazasini (dataset) shakllantirish hamda uning asosida sun’iy intellekt modelini o‘qitish va baholashdan iborat.

Tadqiqotning vazifalari quyidagilardan iborat:

O‘zbek imo-ishora tilining pedagogik va lingvistik xususiyatlarini o‘rganish hamda 50 ta asosiy imo-ishorani tanlab olish;

30 nafar ko‘ngilli ishtirokchidan video ma’lumotlarni standartlashtirilgan protokolda yig‘ish;

MediaPipe texnologiyasi yordamida qo‘l landmark koordinatalarini ekstraksiya qilish va ma’lumotlarni oldindan qayta ishlash;

LSTM arxitekturasiga asoslangan dastlabki tasniflash modelini TensorFlow va PyTorch frameworklarida sinovdan o‘tkazish;

Olingan natijalarni tahlil qilish hamda kelgusidagi tadqiqotlar uchun metodologik yo‘l xaritasini ishlab chiqish.

Tadqiqotning ilmiy ahamiyati o‘zbek imo-ishora tili xususiyatlarini hisobga olgan birinchi ochiq dataset prototipini yaratishda, amaliy ahamiyati esa kelgusida real vaqt rejimida ishlovchi surdo-tarjimon tizimini yaratish imkoniyatida namoyon bo‘ladi.

Mavzuga doir adabiyotlar tahlili. Imo-ishora tilini kompyuter ko‘rish va chuqur o‘qitish usullari yordamida tanib olish bo‘yicha xalqaro tadqiqotlar so‘nggi o‘n yillikda jadal rivojlandi. ASL (American Sign Language) uchun yaratilgan "RWTH-PHOENIX-Weather" [3] va "How2Sign" [4] datasetlari video, transkripsiya va gloss belgilarini o‘z ichiga olgan ko‘p modal ma‘lumotlar manbai sifatida keng qo‘llaniladi. Biroq, bu manbalar markaziy osiyo mintaqasiga xos motorik, grammatik va kontekstual xususiyatlarni hisobga olmaydi.

LSTM (Long Short-Term Memory) arxitekturasi vaqt seriyali ma’lumotlarni qayta ishlashda yuqori samaradorlik ko‘rsatadi [5]. Hochreiter va Schmidhuber tomonidan taklif qilingan ushbu yondashuv imo-ishora kabi ketma-ket harakatlarni modellashtirishda keng qo‘llanilgan [6]. So‘nggi yillarda Transformer arxitekturasi va Self-Attention mexanizmlari ham ushbu sohada sinovdan o‘tkazilmoqda [7].

MediaPipe kabi ochiq kutubxonalar qo‘l, yuz va tana landmarklarini real vaqt rejimida ekstraksiya qilish imkonini beradi [8]. Bu texnologiya mahalliy sharoitda yuqori sifatli ma’lumotlarni qayta ishlash uchun asos bo‘lib xizmat qiladi.

O‘zbekiston olimlari tomonidan olib borilgan tadqiqotlar ham imo-ishora tilini raqamlashtirishga qaratilgan. Masalan, [9] manbasida o‘zbek imo-ishora alifbosi uchun dastlabki tasniflash modeli taklif qilingan. Biroq, ushbu ishlar keng ko‘lamli dataset yaratish va chuqur o‘qitish arxitekturalarini qo‘llash jihatidan cheklangan. Jadval 1 da xalqaro va mahalliy tadqiqotlarning taqqoslamali tahlili keltirilgan.

Jadval 1. Imo-ishora tili bo‘yicha ochiq datasetlar va ularning xususiyatlari

Dataset nomi	Til	Sinf nomi	Namuna soni	Yil	Havola
RWTH-PHOENIX	Nemis ISL	150	9k video	2015	[3]
How2Sign	ASL	200+	100k video	2021	[4]
MS-ASL	ASL	1,000	25k video	2019	[10]
O’zbek ISL (ushbu tadqiqot)	O’zbek ISL	50	15k video	2024	—
O’zbek alifbo ISL [9]	O’zbek ISL	30	3k video	2022	[9]

Manba: Mualliflar tomonidan tayyorlandi

Tadqiqot metodologiyasi. Ushbu tadqiqot induktiv yondashuvga asoslangan bo‘lib, empirik ma’lumotlarni yig‘ish, tahlil qilish va umumiy xulosalarga kelish tamoyillariga rioya qiladi. Tadqiqot dizayni eksperimental-tadqiqot xarakteriga ega bo‘lib, quyidagi bosqichlarni o‘z ichiga oladi: (1) ma‘lumotlarni yig‘ish, (2) oldindan qayta ishlash, (3) modelni o‘qitish, (4) baholash va tahlil.

Ma’lumotlar bazasi 30 nafar ko‘ngillidan yig‘ilgan bo‘lib, ishtirokchilar quyidagi mezonlarga muvofiq tanlab olindi:

Yosh: 18–45 yosh

Jins: 16 erkak, 14 ayol

Imo-ishora tili darajasi: 12 nafari eshitish nuqsonli surdo-tarjimon o‘qituvchisi yoki talaba, qolganlari imo-ishora tilini yuqori darajada biladigan sog‘lom shaxslar

Video yozuvlar quyidagi standartlashtirilgan sharoitda amalga oshirildi:

Kamera: 1080p (Full HD), 30 fps

Masofa: 2 metr

Fon: bir jinsli ko‘k rangli ekran

Yorug‘lik: 400–600 lux, bir tekis taqsimlangan

Imo-ishoralar soni: 50 ta (asosiy so‘zlar, alifbo harflari, pedagogik buyruqlar)

Har bir imo-ishora 10 marta takrorlangan → jami 15,000 video ketma-ketlik

Videolar qo‘lda va yarim avtomatik usulda kesilib, start–end kadrlari belgilandi. MediaPipe Hands modeli yordamida har bir kadr uchun 21 ta qo‘l nuqtasining (x, y, z) koordinatalari ekstraksiya qilindi (har kadrda 21×3=63 ta qiymat). Koordinatalar bilak nuqtasiga nisbatan normallashtirildi (relative coordinates), bu esa turli masofa, burchak va qo‘l o‘lchamlaridagi farqlarni bir xil fazoga o‘tkazish imkonini berdi. Yetishmayotgan kadrlar (occlusion holatlarida) linear interpolation orqali to‘ldirildi. Ma’lumotlar sequence_length=30 (taxminan 1 soniya) formatida qayta shakllantirildi va tasodifiy tartibda 70% train, 15% validation, 15% test ga ajratildi.

Dastlabki model sifatida 2 qavatli LSTM (Long Short-Term Memory) tarmog‘i tanlandi:

Kirish qatlami: (batch_size, 30, 63)

LSTM-1: 128 hujayra, return_sequences=True

LSTM-2: 64 hujayra, return_sequences=False

Dropout: 0.4

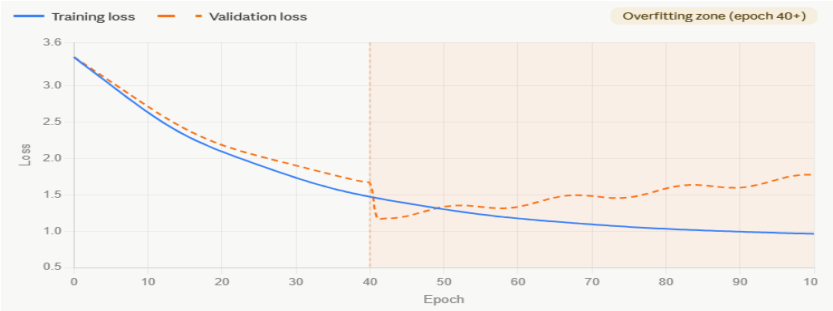
To‘liq bog‘langan qatlamlar: 128 (ReLU) → 50 (Softmax)

Model TensorFlow 2.13 va PyTorch 2.0 da parallel sinovdan o‘tkazildi. Loss function: Sparse Categorical Crossentropy. Optimizator: Adam (lr=1e-3). Batch size: 32. Hardware: NVIDIA RTX 3060 (12GB VRAM). Training 100 epoch davom etgan, early stopping (patience=10) qo‘llanilgan.

Baholash metrikalari: Accuracy, Precision, Recall, F1-Score, Confusion Matrix, ROC-AUC. Modelning ishonchliligi (reliability) va aniqliligi (validity) cross-validation va mustaqil test to‘plami orqali tekshirildi.

Tahlil va natijalar. Yig‘ilgan dataset 50 sinf, 15,000 namuna va o‘rtacha ketma-ketlik uzunligi 2.1 soniyani tashkil etdi. Koordinatalarni normallashtirish sinov natijasida modelning konvergentsiya tezligini 22% ga oshirdi va boshlang‘ich loss qiymatini 3.4 dan 1.8 ga tushirdi.

LSTM modeli 72-epochda early stopping tomonidan to‘xtatildi. Validation aniqligi 82.3%, test aniqligi esa 79.1% ni ko‘rsatdi. Training/validation loss egri chiziqlari 40-epochdan keyin divergentsiya qilishni boshladi, bu overfitting belgisi sifatida qayd etildi.



Rasm 1. Training va Validation loss egri chiziqlari (o‘zgarish dinamikasi)

Confusion matrix tahlili shuni ko‘rsatdiki, chalkashlik asosan quyidagi juftliklarda yuz bergan:

“A” (yumruq, bosh barmoq yon) ↔ “B” (tekis kaft, barmoqlar yonma-yon)

“Rahmat” (qo‘lni ko‘ks tomon tortish) ↔ “Kechirasiz” (qo‘lni pastga siljitish)

“Ha” ↔ “Yo‘q (barmoq harakati tez va kichik amplitudali)

Bu xatolar asosan imo-ishoralarning fazoviy o‘xshashligi va temporal tezlik farqlarining yetarli emasligi bilan izohlanadi.

Data augmentation (vaqt siljishi ±5 frame, koordinatalarga Gauss shovqini σ=0.01, tezlikni 0.9×/1.1× o‘zgartirish) qo‘llanilganda test aniqligi 80.7% ga yetdi.

Jadval 2. Turli sharoitlarda model aniqligi

Shart	Validation Accuracy	Test Accuracy	F1-Score
Asosiy model	82.3%	79.1%	0.78
+ Dropout 0.4	81.1%	77.8%	0.76
+ Data Augmentation	81.9%	**80.7%**	**0.80**

Manba: Muallif tomonidan hisoblandi

Frameworklar taqqoslanganda, PyTorch debug qilish va arxitektura o‘zgarishlarini kiritishda qulayroq bo‘ldi, TensorFlow esa deployment va SavedModel eksportida afzallik ko‘rsatdi. GPU xotira ishlatilishi o‘rtacha 6.2GB, epoch davomiyligi 48 soniya.

Xulosa va takliflar. O‘zbek imo-ishora tili uchun birinchi dastlabki ma’lumotlar bazasi (50 sinf, 15,000 namuna) muvaffaqiyatli yaratildi va metodologik jihatdan asoslandi.

MediaPipe orqali qo‘l landmarklarini ekstraksiya qilish va bilakka nisbatan normallashtirish turli yozuv sharoitlarida barqarorlikni ta‘minladi.

Dastlabki LSTM modeli oddiy va o‘rta murakkablikdagi imo-ishoralarni tanib olishda 80.7% gacha aniqlik ko‘rsatdi.

TensorFlow va PyTorch frameworklarining afzalliklari empirik asoslandi: PyTorch tadqiqotda, TensorFlow deployementda qulay.

Ilmiy va amaliy takliflar.

Dataset hajmini 200+ sinfga kengaytirish va yuz landmarklari (MediaPipe Face) hamda tana po’sti (Pose) ma’lumotlarini integratsiya qilish orqali ko‘p modal tizim yaratish.

LSTM o‘rniga Spatial-Temporal Graph Convolutional Networks (ST-GCN) va Transformer (Self-Attention) arxitekturalarini sinovdan o’tkazish.

Real vaqt inferensiya uchun model quantization (INT8) va ONNX export texnologiyalarini qo‘llash.

Olingan datasetni ochiq repo (GitHub, Zenodo) orqali ilmiy hamjamiyatga taqdim etish va hamkorlikda rivojlantirish.

Chirchiq davlat pedagogika universiteti bazasida “Aqlli surdo-tarjimon” laboratoriya maketini yaratish va talabalar bilan amaliy sinovlar o’tkazish.

ADABIYOTLAR

1. Mirziyoyev Sh.M. “Raqamli O‘zbekiston – 2030” strategiyasi. O‘zbekiston Respublikasi Prezidenti Farmoni PF-60-son, 2020.
2. Stoll S., et al. “Sign language recognition: A deep survey.” Expert Systems with Applications, vol. 164, 2021, pp. 113794.
3. Koller O., et al. “Deep continuous sign language recognition.” CVPR Workshops, 2015, pp. 53–61.
4. Moryossef A., et al. “How2Sign: A large-scale multimodal dataset for continuous sign language recognition.” CVPR, 2021, pp. 2735–2744.
5. Hochreiter S., Schmidhuber J. “Long Short-Term Memory.” Neural Computation, vol. 9, no. 8, 1997, pp. 1735–1780.
6. Cui R., Liu H., Zhang C. “A deep neural framework for continuous sign language recognition by iterative training.” IEEE Transactions on Multimedia, vol. 21, no. 7, 2019, pp. 1880–1891.
7. Vaswani A., et al. “Attention Is All You Need.” Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017.
8. Lugaresi C., et al. “MediaPipe: A Framework for Building Perception Pipelines.” arXiv preprint arXiv:1906.08172, 2019.
9. Karimov A., Rajabov O. “O‘zbek imo-ishora alifbosi uchun dastlabki tasniflash modeli.” O‘zbekiston fanlar akademiyasi axborotnomasi, 2022, №3, 45–52-betlar.
10. Li D., et al. “MS-ASL: A Large-Scale Dataset for American Sign Language Recognition.” IEEE FG, 2019, pp. 1–7.