ISSN: 3030-3346

РАСПОЗНАВАНИЕ ЖЕСТОВЫХ ЯЗЫКОВ МЕТОДАМИ КОМПЬЮТЕРНОГО ЗРЕНИЯ: ОБЗОР И МЕСТО УЗБЕКСКОГО ЖЕСТОВОГО ЯЗЫКА

Онаркулов Максад Каримбердиевич[1], Юсупов Мирсаид Абдулазиз угли[2], Умиржонов Лазизжон Акмалжон угли[3]

[1]Доктор философии (PhD) по физико-математическим наукам, доцент, декан физико-математического факультета, Ферганский государственный университет, г. Фергана, Республика Узбекистан. E-mail: maxmaqsad@gmail.com. ORCID: 0000-0002-3650-4847

[2]Преподаватель кафедры прикладной математики и информатики, Ферганский государственный университет, г. Фергана, Республика Узбекистан. E-mail: mirsaidbeky@gmail.com. ORCID: 0009-0001-1989-0088

[3]Независимый исследователь кафедры прикладной математики и информатики, Ферганский государственный университет, г. Фергана, Республика Узбекистан. E-mail: lazizbekq8@gmail.com. ORCID: 0009-0007-0326-3921

K E Y W O R D S

распознавание жестового языка, компьютерное зрение, узбекский жестовый язык, ключевые точки, перенос обучения, малоресурсные языки, картирующий обзор

A B S T R A C T

Мы систематизировали исследования 2017–2026 гг. по автоматическому распознаванию жестовых языков методами компьютерного зрения и попытались определить, какое место в этой области занимает узбекский жестовый язык (УзЖЯ). Поводом послужили постановление Президента Республики Узбекистан № ПП-407 о развитии УзЖЯ и острая нехватка сурдопереводчиков в стране. Поиск вёлся в научных базах данных, архивах препринтов и репозиториях на английском, русском и узбекском языках, а отбор и описание публикаций выполнены по рекомендациям PRISMA-ScR для картирующих обзоров. Из 220 найденных записей в анализ вошли 196. Мы рассмотрели наборы данных, способы представления жеста, архитектуры моделей, перенос обучения, облегчённые модели для мобильных устройств и протоколы оценки. Открытые данные сосредоточены вокруг немногих языков, прежде всего американского и русского; жестовых языков Центральной Азии в многоязычных корпусах нет совсем. По УзЖЯ нашлось 13 технических публикаций 2023–2026 гг. и один открытый программный репозиторий. Ни в одной работе не опубликован набор данных и не проведена проверка на исполнителях, не участвовавших в обучении, так что их результаты между собой несравнимы. Технические работы по УзЖЯ здесь впервые сопоставлены по единым критериям. Кроме того, по явному правилу оценено состояние 26 направлений в мире, в регионе и для УзЖЯ, и 24 из них отнесены к исследовательским пробелам. Для УзЖЯ первоочередными мы считаем четыре задачи: адаптацию моделей, предобученных на корпусах русского жестового языка, с разделением общих и собственных жестов; распознавание с отклонением жестов вне словаря; подстройку модели под нового пользователя без разметки; совместное распознавание кириллического и латинского дактиля.

ВВЕДЕНИЕ

По оценке Всемирной организации здравоохранения, реабилитация в связи с потерей слуха нужна более чем 430 млн человек, а к 2050 г. таких людей может стать свыше 700 млн [1]. В Узбекистане на учёте Общества глухих состоит более 22 тыс. человек и около 8 000 детей с инвалидностью по слуху. В столице на одного квалифицированного переводчика жестового языка приходится около 800 человек с нарушением слуха, а первый выпуск дипломированных переводчиков, по данным на сентябрь 2025 г., ожидался только в 2026 г. [2]. Сколько людей пользуется жестовым языком, официально не установлено: в предварительных итогах переписи такие ответы попали в категорию «Другие», а окончательные данные Национальный комитет по статистике обещает опубликовать только в 2027 г. [3].

(cc) BY

Постановление Президента Республики Узбекистан от 20.10.2022 № ПП-407 [4] предусматривает разработку дактильной азбуки узбекского жестового языка (УзЖЯ) на латинской графике и его лингвистического словаря. Кроме того, не менее 5 % недельного эфира государственных телеканалов должно транслироваться с сурдопереводом, скрытыми субтитрами или бегущей строкой. Словарь к установленному сроку, 1 июня 2024 г., подготовлен не был, и онлайн-словарь объёмом более 2 000 слов с видеопримерами создаёт общественное объединение «Sharoit Plyus» [2, 5].

По оценкам специалистов, жестовый язык Узбекистана мало отличается от русского жестового языка (РЖЯ), и в быту глухие широко используют русский дактильный алфавит на кириллице, тогда как ПП-407 вводит латинский [4, 6, 7]. Близость подтверждается и генеалогическими классификациями: в базе Glottolog казахско-русский и таджикский жестовые языки описаны как разновидности РЖЯ, а отдельной записи для УзЖЯ там нет [8]. Лексикостатистических исследований соотношения УзЖЯ и РЖЯ мы не нашли. Между тем для РЖЯ опубликованы крупные открытые корпуса [9, 10]. Когда у малоресурсного языка есть близкий родственник, хорошо обеспеченный данными, задача распознавания ставится иначе, чем для изолированного языка. На это обстоятельство мы опираемся во всей работе.

Обзоры последних лет [11–13] описывают переход от ручного конструирования признаков и специальных датчиков к глубоким нейронным сетям, которые работают с обычным видео. Появились наборы из десятков тысяч видеозаписей [14, 15] и многоязычный корпус YouTube-SL-25 объёмом 3 207 ч [16], модели на последовательностях ключевых точек тела и кистей [17, 18], методы самообучаемого предобучения [19, 20]. Но распределены эти достижения неравномерно. В обзоре Ni и соавторов [21], охватывающем 120 наборов данных для 35 жестовых языков, из языков Центральной Азии представлен только казахско-русский (корпус K-RSL), а узбекского нет; в YouTube-SL-25 языков региона нет совсем. Обзоры Berzhanova и соавт. [22] и Alishzade и Abdullayeva [23] посвящены видеораспознаванию и малоресурсным жестовым языкам. Во втором из них УзЖЯ отнесён к языкам начального уровня технологического развития, однако конкретные работы по УзЖЯ, их данные и протоколы оценки не разбираются. Русскоязычные обзоры [12, 13] ориентированы на РЖЯ.

Цель работы — систематизировать исследования 2017–2026 гг. по распознаванию жестовых языков методами компьютерного зрения и определить место УзЖЯ в этой области. Для этого мы отобрали публикации по рекомендациям PRISMA-ScR, проанализировали наборы данных, методы и протоколы оценки, сравнили технические работы по УзЖЯ по единым критериям и выявили исследовательские пробелы. Наборам данных жестовых языков была посвящена наша предыдущая статья [24]; здесь в центре внимания методы распознавания и протоколы оценки. Насколько нам известно, систематизированного обзора технических работ по УзЖЯ до сих пор не публиковалось.

МАТЕРИАЛЫ И МЕТОДЫ

Обзор выполнен как картирующий (scoping review) и оформлен по рекомендациям PRISMA-ScR [25] с опорой на PRISMA 2020 [26]. Такой формат выбран потому, что нам важно описать состояние области и найти пробелы, а не обобщить эффект одного метода; к тому же англоязычную литературу мы охватывали выборочно. Поиск проводился в сентябре 2026 г. по Crossref, OpenAlex, arXiv, ACL Anthology, IEEE Xplore (по метаданным), Zenodo и КиберЛенинке, а также по репозиториям университетов и базе авторефератов ОАК Республики Узбекистан. Запросы формулировались на английском, русском и узбекском языках (латиница и кириллица), например «sign language recognition», «распознавание жестового языка», «imo-ishora tilini tanib olish» и их варианты.

Включались работы по распознаванию отдельных жестов, непрерывной жестовой речи и дактиля, по переводу с жестового языка

(cc) BY

на звучащий, описания наборов данных и обзоры. Исключались работы по синтезу жестов, переводу текста в глоссы и смежным задачам без распознавания. Публикации по РЖЯ, казахскому (КЖЯ) и казахско-русскому (КРЖЯ) жестовым языкам и по УзЖЯ мы старались собрать полностью; по остальным языкам отбирались ключевые англоязычные работы. Для каждой записи отмечалось, по чему она оценена: по полному тексту, аннотации или метаданным. Без полного текста оценено 68 работ. Каталог включённых работ авторы предоставят по запросу.

Всего найдено 220 записей: 206 в базах данных и 14 в других источниках (уточняющий поиск и списки литературы обзоров). Исключены 24 записи, по 8 в каждой группе: синтез жестов, перевод текста в глоссы, смежные темы. В анализ вошли 196 работ, из них 11 из других источников (рис. 1).



Рисунок 1 – Схема отбора публикаций (по PRISMA-ScR)

Англоязычные работы отбирались выборочно, поэтому динамику публикаций мы показываем только для сплошь учтённых работ по России, Казахстану и Узбекистану (рис. 2). Их число выросло с 1–2 в год в 2017–2018 гг. до 24 в 2024 г. и 18 в 2025 г. Основной прирост дали исследователи Казахстана: 25 работ за 2024–2025 гг. Российские работы распределены по годам ровнее, от 1 до 7 в год. Первая техническая публикация по УзЖЯ вышла в 2023 г., а восемь из 13 — в 2024 г.

(cc) BY

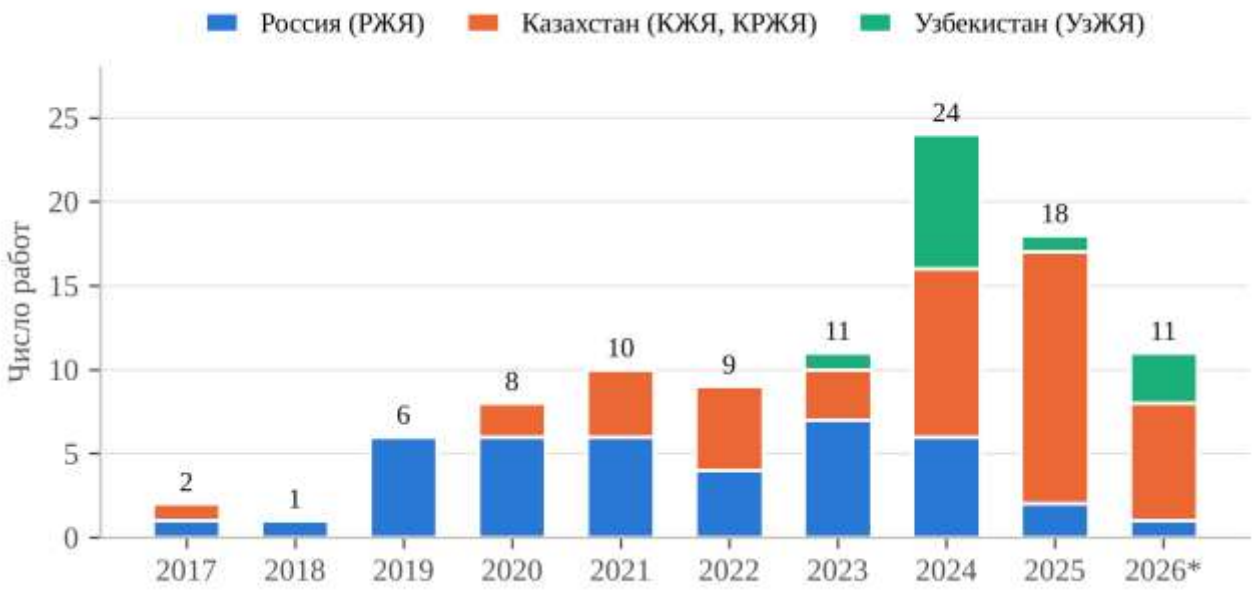


Рисунок 2 – Число работ по годам: Россия, Казахстан, Узбекистан (сплошной учёт, n = 100; без англоязычной выборки, уточняющего поиска и программного репозитория; 2026 г. — январь–сентябрь)

РЕЗУЛЬТАТЫ

Наборы данных

Подробно наборы данных жестовых языков разобраны в нашей предыдущей работе [24], поэтому в табл. 1 собраны только те, что важны для работ по УзЖЯ. Самый крупный открытый корпус отдельных жестов РЖЯ, Logos [10], содержит 199 668 видеозаписей 2 863 глосс от 381 исполнителя, почти в 10 раз больше, чем Slovo [9]. Кириллический дактиль РЖЯ представлен набором Bukva [27]. В нём 33 буквы русского дактильного алфавита, того самого, которым глухие в Узбекистане пользуются в быту [7]; букв Ў, Қ, Ғ и Ҳ узбекской кириллицы в нём нет. Для казахско-русского жестового языка созданы корпуса K-RSL [28] и FluentSigners-50 [29]; второй с самого начала задуман для оценки на новых исполнителях. Собственные наборы есть также у азербайджанского (AzSLD) и турецкого (AUTSL) жестовых языков [15, 30].

Немаловажен и правовой статус данных. Logos распространяется по лицензии Apache-2.0, Slovo и Bukva — по лицензиям на основе CC BY-SA 4.0 [9, 10, 27]. Видеоматериалы онлайн-словаря УзЖЯ защищены авторским правом [5], и обучать на них модели можно только с согласия правообладателя. Открытых наборов данных УзЖЯ на момент подготовки статьи мы не обнаружили.

Таблица 1

Основные наборы данных для распознавания жестовых языков (по данным авторов наборов)

Набор	Жестовый язык	Объём	Значение для УзЖЯ
Logos [10]	РЖЯ	199 668 видео, 2 863 глоссы, 381 исполнитель, 221,4 ч; Apache-2.0	Основной корпус для предобучения
Slovo [9]	РЖЯ	20 400 видео, 1 000 классов, 194 исполнителя	Предобучение, базовые модели
Bukva [27]	РЖЯ, дактиль (кириллица)	3 757 видео, 33 буквы, 155 исполнителей	Дактиль, используемый в быту (без букв Ў, Қ, Ғ, Ҳ)

(cc) BY

K-RSL; FluentSigners-50 [28, 29]	КРЖЯ	Корпус для лингвистики и распознавания; эталон оценки на новых исполнителях	Образец протокола оценки
AzSLD [30]	Азербайджанский ЖЯ	≈30 000 видео, съёмка двумя камерами	Пример национального набора тюркоязычной страны
AUTSL [15]	Турецкий ЖЯ	38 336 видео, 226 жестов, 43 исполнителя	Эталон оценки на новых исполнителях
WLASL [14]	Американский ЖЯ	≈21 тыс. видео, 2 000 слов, 119 исполнителей	Крупный эталон пословного распознавания
YouTube-SL-25 [16]	Более 25 ЖЯ	3 207 ч, из них РЖЯ — 60 ч; ЖЯ Центральной Азии нет	Неразмеченные видео РЖЯ для самообучения

Методы распознавания

Подходы различаются прежде всего тем, как представлен жест. Модели, работающие с цветным видео (трёхмерные свёрточные сети, видеотрансформеры), извлекают признаки прямо из пикселей и требуют больших вычислительных ресурсов. Есть и другой путь: работать с последовательностями ключевых точек тела, кистей и лица, которые дают оценщики позы, чаще всего MediaPipe [17]. Видеозапись жеста из T кадров тогда записывается как последовательность

$$X = (x_1, x_2, \dots, x_T), \qquad x_t \in \mathbb{R}^{K\times 3} \qquad (1)$$

где x_t — координаты K ключевых точек в кадре t (две координаты на плоскости изображения и относительная глубина). У MediaPipe Holistic K = 543, но на практике обычно берут только точки кистей и верхней части тела. Такое представление компактно, мало зависит от фона и подходит для мобильных устройств. Если же объединить скелетные и визуальные признаки в многопотоковой схеме, точность растёт [18]. На итоговое качество влияет и выбор оценщика позы: при сравнении восьми оценщиков на задаче перевода SDPose и Sapiens дали BLEU около 11,5, а MediaPipe — около 10 [31].

Архитектуры прошли путь от сочетаний свёрточных и рекуррентных сетей к трансформерам [12, 13]. SignBERT+ предобучает представление кисти на неразмеченных последовательностях ключевых точек [19], а Uni-Sign решает в одной модели задачи распознавания отдельных жестов, непрерывной речи и перевода [20].

Для малоресурсных языков решающим становится перенос обучения, то есть использование моделей, предобученных на других жестовых языках. В проекте OpenHands [32] опубликованы предобученные на позах модели для нескольких жестовых языков. Авторы корпуса Logos [10] показали, что предобучение на нём с последующим совместным обучением повышает точность и на сторонних наборах: 97,92 % на Logos, 97,81 % на AUTSL и 66,82 % на WLASL. Bilge и соавт. [33] поставили задачу межъязыкового распознавания по нескольким примерам и предложили для неё три многоязычных тестовых набора. Kindiroglu и соавт. [34] исследовали перенос между двумя наборами турецкого жестового языка, в том числе при частичном совпадении классов. Ryumin и соавт. [35] использовали для межъязыкового переноса визуальную информацию. Работ, где модель адаптируют к близкородственному жестовому языку с частично общим словарём, как в паре РЖЯ — УзЖЯ, в нашей выборке нет. Работа [34] ближе всего к этой постановке, но в ней речь идёт о разных наборах одного языка, а не о родственных языках.

О пригодности модели на практике судят по её точности на людях, которых не было в обучающей выборке. Такое разбиение заложено в AUTSL и FluentSigners-50 [15, 29]. Если же делить записи на обучающие и тестовые случайно, модель запоминает манеру

(cc) BY

конкретных исполнителей и точность получается завышенной. Методы, отклоняющие жесты вне словаря (распознавание при открытом множестве классов), разработаны для распознавания действий по скелету [36], а подстройка под нового пользователя без разметки на этапе применения — для перевода жестового языка [37]. Для распознавания отдельных жестов подобных работ мы не нашли.

Для мобильных систем предложены облегчённые модели на ключевых точках, например с дистилляцией знаний от трёхмерных моделей к одномерным [38]. Большие мультимодальные модели специализированные пока не заменяют. На WLASL300 Gemini-2.5-pro правильно назвала 25,45 % жестов с учётом синонимов (32,28 %, если модели дан список глосс), тогда как специализированный классификатор DSLNet достигает 89,97 %; открытые мультимодальные модели отстают ещё сильнее [39].

Исследования в регионе и по узбекскому жестовому языку

В регионе дальше всех продвинулась санкт-петербургская школа исследований РЖЯ. В диссертации Д. А. Рюмина [40] разработаны модели и методы автоматического распознавания элементов РЖЯ для человеко-машинного взаимодействия; он же участвовал в подготовке аналитических обзоров [12, 13]. После публикации корпусов Slovo, Bukva и Logos [9, 10, 27] РЖЯ вошёл в число жестовых языков, лучше всего обеспеченных данными.

Казахстанские исследователи создали корпуса K-RSL и FluentSigners-50 [28, 29]. С. Б. Муханов в диссертации на степень PhD разработал распознавание 42 букв казахского дактиля в реальном времени [41]. Zhassuzak и соавт. [42] предложили систему двунаправленного перевода с учётом просодии; точность 92 % получена на 12 предложениях, а оценка на новых исполнителях не описана. Для Кыргызстана, Таджикистана и Туркменистана работ по распознаванию жестов методами компьютерного зрения мы не нашли.

В 2026 г. в Узбекистане зарегистрированы темы диссертаций по распознаванию УзЖЯ, машинному переводу узбекской речи на жестовый язык и созданию наборов данных УзЖЯ [43]. Мы выявили 13 технических публикаций по УзЖЯ за 2023–2026 гг. и один открытый программный репозиторий (табл. 2); на рис. 1 они учтены вместе, как 14 записей.

Первая по времени работа Kayumov и Kayumova [44] обсуждает, насколько алгоритмы машинного обучения применимы к УзЖЯ, но численных результатов не приводит. Те же авторы предложили распознавание латинского дактиля по угловым признакам суставов пальцев и запястья [45], классификатор жестов на свёрточных и рекуррентных сетях с переносом обучения [46] и модель распознавания слов на основе свёрточной сети ResNet-50 [47]. Akhmedov и соавт. [48] применили к дактилю детектор семейства YOLO и привели метрики precision, recall и AP; алгоритм распознавания букв описан также у Urinov и Umarov [49]. Jovliyeva [50] получила 99,11 % для статических букв (MediaPipe Hands и случайный лес) и 65,96 % для динамических (LSTM) на 1 523 кадрах и 468 последовательностях. Для работ [51, 52] доступны только аннотации. Musaev и соавт. [53] и Juraev и соавт. [54] занимаются переводом, причём первая работа — в направлении «речь → жест».

Самый большой словарь среди опубликованных работ, более 500 жестов, собрали Ergashev и соавт. [55]. В работе [56] эти авторы сравнили две модели на ключевых точках: Squeezeformer показала на закрытом наборе точность 91 % (CNN-LSTM — 87 %) при задержке 0,58 с. Сами авторы отмечают, что разбиение данных не было независимым от исполнителя. Похожая архитектура раньше применялась в открытом решении соревнования по распознаванию дактиля американского жестового языка [57]. На основе этой работы создана мобильная обучающая платформа SIGNLY.UZ [55].

(cc) BY

Таблица 2

Технические работы по распознаванию и переводу УзЖЯ (2023–2026 гг.)

№	Работа	Задача и данные	Метод и результат (по данным авторов)	Открытые данные / оценка на новых исполнителях
1	Kayumov, Kayumova, 2023 [44]	Распознавание УзЖЯ	Обзор алгоритмов машинного обучения; численных результатов нет	— / —
2	Akhmedov и др., 2024 [48]	Дактиль по изображениям	Детектор семейства YOLO; precision, recall, AP	Нет / не описана
3	Urinov, Umarov, 2024 [49]	Буквы УзЖЯ	Алгоритм распознавания букв	Нет данных
4	Kayumov, Kayumova, 2024 [45]	Латинский дактиль	Угловые признаки суставов пальцев и запястья, свёрточная сеть	Нет данных
5	Kayumov и др., 2024 [46]	Классификатор жестов	CNN и RNN с переносом обучения	Нет / не описана
6	Kayumov и др., 2024 [47]	Слова УзЖЯ, кадры видео	Свёрточная сеть ResNet-50	Нет данных
7	Saroliya и др., 2024 [51]	Распознавание жестов	Машинное обучение	Нет данных
8	Sabitjanov и др., 2024 [52]	Распознавание жестов	Машинное обучение	Нет данных
9	Musaev и др., 2024 [53]	Перевод речи в жесты	Обработка текста для системы «речь → жест»	Распознавание жестов не рассматривается
10	Juraev и др., 2025 [54]	Перевод УзЖЯ	MediaPipe Holistic, распознавание речи, языковая модель BERT	Нет данных
11	Jovliyeva, 2026 [50]	Дактиль: 1 523 кадра, 468 последовательностей	MediaPipe Hands + случайный лес — 99,11 % (статика); LSTM — 65,96 % (динамика)	Нет / не описана
12	Ergashev и др., 2026 [56]	Отдельные жесты, закрытый набор ключевых точек	Squeezeformer — 91 %, CNN-LSTM — 87 %; задержка 0,58 и 0,72 с	Нет / нет (указано авторами)
13	Ergashev и др., 2026 [55]	Мобильная обучающая платформа; набор более 500 жестов	Flutter, TensorFlow Lite, офлайн-распознавание части словаря; пилот на 10 участниках	Нет / нет данных
—	Репозиторий surdo_Uzsl [58]	Мобильный переводчик, Telegram-бот	MediaPipe Holistic, LSTM или трансформер; YOLOv8 для дактиля	Код открыт; метрик нет

Ни в одной из этих работ набор данных не опубликован, поэтому результаты несравнимы. Оценки на новых исполнителях нет нигде; все системы распознают жесты из закрытого словаря, а кириллический и латинский дактиль рассматриваются по отдельности. Задержку измеряли только в [56], энергопотребление на мобильных устройствах не оценивал никто.

Сейчас мы разрабатываем систему UZSign.AI для распознавания отдельных

(cc) BY

жестов УзЖЯ по ключевым точкам; её словарь включает 517 лексических единиц, а результаты будут опубликованы отдельно. Больше всего при сборе данных нам мешает нехватка исполнителей.

ОБСУЖДЕНИЕ

Состояние исследований мы оценили по 26 направлениям на трёх уровнях: в мире, в регионе (Центральная Азия, страны СНГ, Турция) и для УзЖЯ (рис. 3). Каждому уровню присваивался один из статусов: «развито», «частично», «мало» (единичные работы) или «нет». Пробел считался глобальным, если в мире работ по направлению нет или они единичны; региональным, если в мире работы есть, а в регионе их нет или они единичны; локальным, если работы есть в регионе, но не для УзЖЯ. Если для УзЖЯ направление развито хотя бы частично, пробелом оно не считалось. Отсутствие работ мы признавали только тогда, когда не менее трёх запросов не менее чем в двух поисковых системах не давали релевантных результатов. По этому правилу 24 направления из 26 оказались пробелами: 8 глобальных, 11 региональных и 5 локальных.

Из глобальных пробелов четыре, на наш взгляд, стоит решать для УзЖЯ в первую очередь: для них хватит открытых корпусов РЖЯ и собственных данных авторов, без массовой новой записи.

1) Адаптация моделей, предобученных на корпусах РЖЯ [9, 10], к УзЖЯ. Историческая связь языков [6–8] делает такую адаптацию обоснованной, но жесты придётся разделить на общие для двух языков и собственные жесты УзЖЯ. Доля последних становится параметром эксперимента.

2) Распознавание с отклонением жестов вне словаря. Словарь любой системы меньше лексики языка: один только онлайн-словарь содержит более 2 000 слов [5]. Проще всего принимать решение с порогом по максимальной апостериорной вероятности [59]:

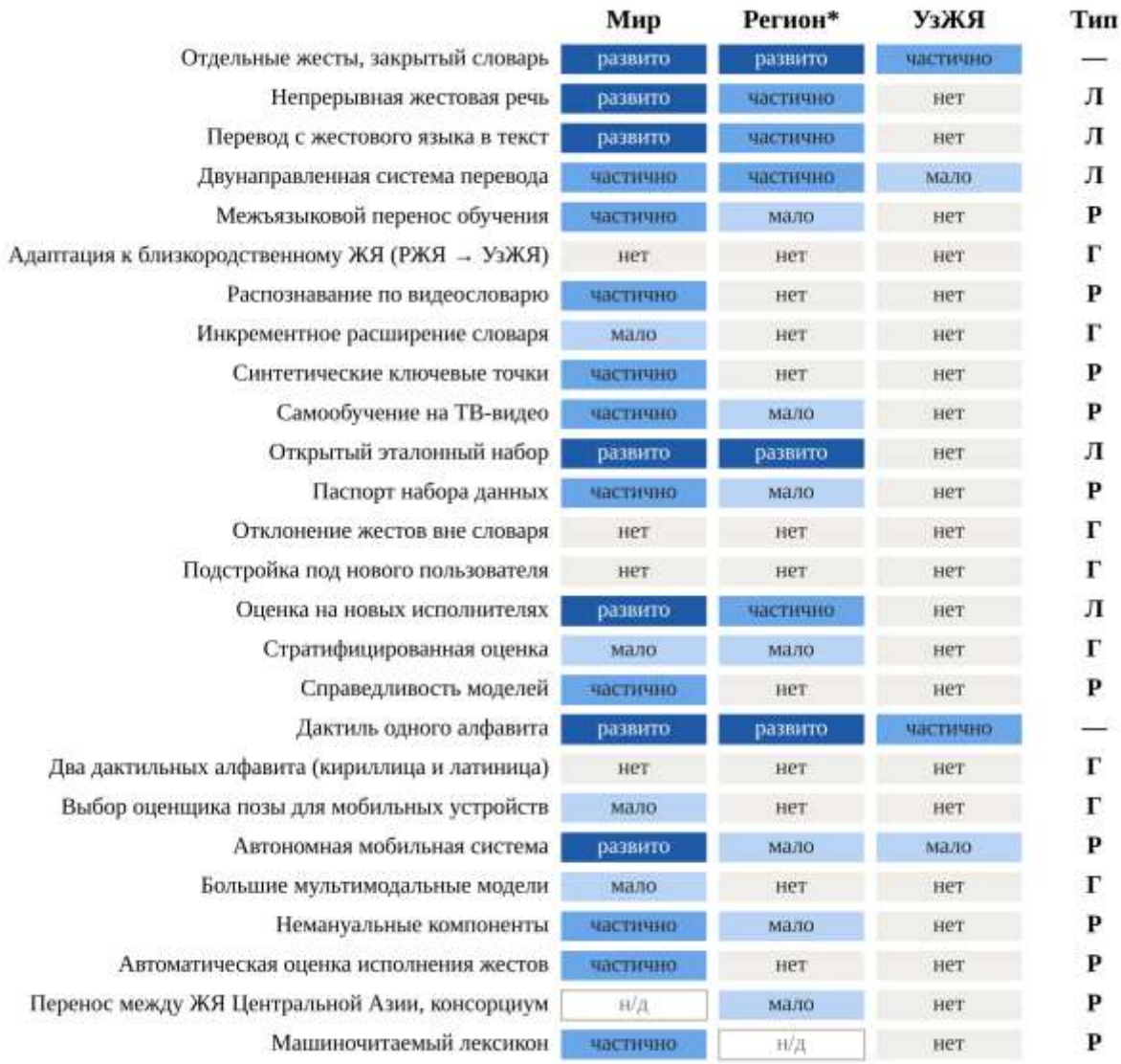
$$\hat{y} = \arg\max_{y \in Y} p(y \mid X) \quad при \quad \max_{y \in Y} p(y \mid X) \geq \tau \qquad (2)$$

где Y — словарь системы; p(y|X) — оценка вероятности класса y для последовательности (1); τ — порог, подбираемый на проверочной выборке. Если условие не выполнено, система отказывается отвечать, а не выдаёт уверенную ошибку на незнакомом жесте или переходном движении. Такой порог служит лишь отправной точкой: известно, что softmax-оценки нейросетей бывают завышены, поэтому в дальнейшем мы планируем сравнить его с более устойчивыми оценками.

3) Подстройка модели под нового пользователя на этапе применения, без разметки. Если исполнителей в обучающих данных немного, в эксплуатации модель неизбежно столкнётся с манерой исполнения, которой при обучении не было.

4) Совместное распознавание кириллического дактиля, которым пользуются глухие, и латинского, введённого ПП-407 [4, 7].

(cc) BY



	Мир	Регион*	УзЖЯ	Тип
Отдельные жесты, закрытый словарь	развито	развито	частично	—
Непрерывная жестовая речь	развито	частично	нет	Л
Перевод с жестового языка в текст	развито	частично	нет	Л
Двунаправленная система перевода	частично	частично	мало	Л
Межъязыковой перенос обучения	частично	мало	нет	Р
Адаптация к близкородственному ЖЯ (РЖЯ → УзЖЯ)	нет	нет	нет	Г
Распознавание по видеословарю	частично	нет	нет	Р
Инкрементное расширение словаря	мало	нет	нет	Г
Синтетические ключевые точки	частично	нет	нет	Р
Самообучение на ТВ-видео	частично	мало	нет	Р
Открытый эталонный набор	развито	развито	нет	Л
Паспорт набора данных	частично	мало	нет	Р
Отклонение жестов вне словаря	нет	нет	нет	Г
Подстройка под нового пользователя	нет	нет	нет	Г
Оценка на новых исполнителях	развито	частично	нет	Л
Стратифицированная оценка	мало	мало	нет	Г
Справедливость моделей	частично	нет	нет	Р
Дактиль одного алфавита	развито	развито	частично	—
Два дактильных алфавита (кириллица и латиница)	нет	нет	нет	Г
Выбор оценщика позы для мобильных устройств	мало	нет	нет	Г
Автономная мобильная система	развито	мало	мало	Р
Большие мультимодальные модели	мало	нет	нет	Г
Немануальные компоненты	частично	мало	нет	Р
Автоматическая оценка исполнения жестов	частично	нет	нет	Р
Перенос между ЖЯ Центральной Азии, консорциум	н/д	мало	нет	Р
Машиночитаемый лексикон	частично	н/д	нет	Р

*Рисунок 3 – Карта исследовательских пробелов (Г — глобальный, Р — региональный, Л — локальный пробел, «—» — не пробел; н/д — нет данных; * Центральная Азия, страны СНГ, Турция; оценка авторов по источникам обзора)*

Для решения этих задач нужна исследовательская инфраструктура: открытый эталонный набор УзЖЯ, единый протокол оценки на новых исполнителях с разбивкой результатов по региону, ведущей руке, освещённости и типу устройства, паспорт набора данных по образцу [21]. Для мобильных приложений нужен обоснованный выбор оценщика позы [31] и измерение задержки, расхода памяти и энергопотребления. На уровне региона перспективен консорциум тюркских жестовых языков [23]. Создавать эти ресурсы стоит вместе с сообществом глухих и сурдопереводчиками.

У нашего обзора есть ограничения. Отбор и извлечение данных выполнялись без независимого двойного скрининга; 68 работ оценены только по аннотациям и метаданным; англоязычная литература представлена выборкой. Если работ по направлению не нашлось в просмотренных источниках, это ещё не значит, что их нет: вывод о пробеле относится к сентябрю 2026 г. и к перечисленным базам, а препринты последних месяцев могли быть проиндексированы не полностью. Статусы направлений на рис. 3 отражают экспертную оценку авторов.

ЗАКЛЮЧЕНИЕ

1. В картирующий обзор по рекомендациям PRISMA-ScR вошли 196 работ 2017–2026 гг. из 220 найденных записей. В

(cc) BY

мировой практике утвердились модели на ключевых точках, трансформеры и предобучение на крупных корпусах. Открытые данные при этом сосредоточены вокруг нескольких жестовых языков, а языков Центральной Азии в многоязычных корпусах нет.

2. По УзЖЯ найдено 13 технических публикаций и один открытый программный репозиторий. Наборы данных не опубликованы, оценка на новых исполнителях не проводилась, поэтому приведённые значения точности несравнимы.

3. Историческая связь УзЖЯ и РЖЯ и наличие крупных открытых корпусов РЖЯ делают адаптацию предобученных моделей самым реалистичным способом преодолеть нехватку данных, если разделить общие и собственные жесты УзЖЯ.

4. Из 26 направлений 24 отнесены к пробелам: 8 глобальных, 11 региональных и 5 локальных. Первоочередными для УзЖЯ мы считаем адаптацию моделей РЖЯ, распознавание с отклонением жестов вне словаря, подстройку под нового пользователя и совместное распознавание двух дактильных алфавитов. Для всех четырёх задач нужны открытый эталонный набор и единый протокол оценки. Этим задачам будут посвящены наши дальнейшие исследования.

СПИСОК ЛИТЕРАТУРЫ

1. Deafness and hearing loss [Электронный ресурс]: информационный бюллетень / Всемирная организация здравоохранения. 2026. 3 марта. URL: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss (дата обращения: 06.10.2026).
2. Юсупов Д., Ахлидинов М. Как в Узбекистане реализуется право глухих людей на жестовый язык? [Электронный ресурс] // Газета.uz. 2025. 28 сентября. URL: https://www.gazeta.uz/ru/2025/09/28/sign-language/ (дата обращения: 06.10.2026).
3. Окончательных данных о числе носителей узбекского жестового языка придется ждать до 2027 года [Электронный ресурс] // Anhor.uz. 2026. 15 июля. URL: https://anhor.uz/news/zhestovyy-yazyk-perepis-uzbekistan/ (дата обращения: 06.10.2026).
4. О дополнительных мерах по развитию узбекского языка жестов и азбуки Брайля: постановление Президента Республики Узбекистан от 20.10.2022 № ПП-407 [Электронный ресурс] // Национальная база данных законодательства Республики Узбекистан. URL: https://lex.uz/ru/docs/6248469 (дата обращения: 06.10.2026).
5. UzSL — Uzbekistan Sign Language [Электронный ресурс]: онлайн-словарь узбекского жестового языка / ОО инвалидов «SHAROIT PLYUS». Ташкент, 2026. URL: https://imo-ishora.uz/ (дата обращения: 06.10.2026).
6. Ахлиддинов М., Юсупов Д. Быть ли узбекскому жестовому языку? [Электронный ресурс] // Газета.uz. 2020. 23 сентября. URL: https://www.gazeta.uz/ru/2020/09/23/sign-language/ (дата обращения: 06.10.2026).
7. Ахлидинов М., Юсупов Д. Постановление о нас без нас: узбекский жестовый язык хотят развивать без участия его носителей? [Электронный ресурс] // Газета.uz. 2022. 22 апреля. URL: https://www.gazeta.uz/ru/2022/04/22/sign-language/ (дата обращения: 06.10.2026).
8. Hammarström H., Forkel R., Haspelmath M., Bank S. Glottolog: Russian Sign Language family (RSLic) [Электронный ресурс]. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: https://glottolog.org/resource/languoid/id/russ1270 (дата обращения: 07.10.2026).
9. Kapitanov A., Kvanchiani K., Nagaev A., Petrova E. Slovo: Russian Sign Language Dataset // Computer Vision Systems (ICVS 2023). Lecture Notes in Computer Science. Cham: Springer, 2023. P. 63–73. DOI: 10.1007/978-3-031-44137-0_6.
10. Ovodov I., Surovtsev P., Kvanchiani K. [et al.] Logos as a Well-Tempered Pre-train for Sign Language Recognition // Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing.

(cc) BY

Suzhou: ACL, 2025. P. 24340–24353. DOI: 10.18653/v1/2025.emnlp-main.1238.
11. Koller O. Quantitative Survey of the State of the Art in Sign Language Recognition [Электронный ресурс] // arXiv. 2020. URL: https://arxiv.org/abs/2008.09918 (дата обращения: 06.10.2026).
12. Рюмин Д.А., Кагиров И.А., Аксёнов А.А., Карпов А.А. Аналитический обзор моделей и методов автоматического распознавания жестов и жестовых языков // Информационно-управляющие системы. 2021. № 6. С. 10–20. DOI: 10.31799/1684-8853-2021-6-10-20.
13. Иванько Д.В., Рюмин Д.А. Автоматический сурдоперевод: обзор нейросетевых методов распознавания и синтеза звучащей и жестовой речи // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24, № 5. С. 669–686. DOI: 10.17586/2226-1494-2024-24-5-669-686.
14. Li D., Rodriguez C., Yu X., Li H. Word-level Deep Sign Language Recognition from Video: A New Large-scale Dataset and Methods Comparison // 2020 IEEE Winter Conference on Applications of Computer Vision (WACV). 2020. DOI: 10.1109/WACV45572.2020.9093512.
15. Sincan O.M., Keles H.Y. AUTSL: A Large Scale Multi-modal Turkish Sign Language Dataset and Baseline Methods // IEEE Access. 2020. Vol. 8. P. 181340–181355. DOI: 10.1109/ACCESS.2020.3028072.
16. Tanzer G., Zhang B. YouTube-SL-25: A Large-Scale, Open-Domain Multilingual Sign Language Parallel Corpus [Электронный ресурс] // The Thirteenth International Conference on Learning Representations (ICLR 2025). 2025. URL: https://arxiv.org/abs/2407.11144 (дата обращения: 06.10.2026).
17. Lugaresi C., Tang J., Nash H. [et al.] MediaPipe: A Framework for Building Perception Pipelines [Электронный ресурс] // arXiv. 2019. URL: https://arxiv.org/abs/1906.08172 (дата обращения: 06.10.2026).
18. Jiang S., Sun B., Wang L. [et al.] Skeleton Aware Multi-modal Sign Language Recognition // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2021). 2021. P. 3413–3423.
19. Hu H., Zhao W., Zhou W., Li H. SignBERT+: Hand-model-aware Self-supervised Pre-training for Sign Language Understanding // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2023. Vol. 45, No. 9. P. 11221–11239.
20. Li Z., Zhou W., Zhao W. [et al.] Uni-Sign: Toward Unified Sign Language Understanding at Scale [Электронный ресурс] // The Thirteenth International Conference on Learning Representations (ICLR 2025). 2025. URL: https://arxiv.org/abs/2501.15187 (дата обращения: 06.10.2026).
21. Ni Y., Cheng Z.-Q., Li J., Cheng W. Sign-Language Datasets at Scale: A Comprehensive Survey on Resources, Benchmarks, and Annotation Standards // Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics. Vol. 1: Long Papers. San Diego: ACL, 2026. P. 41578–41604. DOI: 10.18653/v1/2026.acl-long.1928.
22. Berzhanova U., Yerimbetova A., Milosz M. [et al.] A Comprehensive Review of Deep Learning Approaches for Video-Based Sign Language Recognition: Datasets, Challenges and Insights // Multimodal Technologies and Interaction. 2026. Vol. 10, No. 6. Art. 58. DOI: 10.3390/mti10060058.
23. Alishzade N., Abdullayeva G. Sign Language Recognition and Translation for Low-Resource Languages: Challenges and Pathways Forward [Электронный ресурс] // arXiv. 2026. URL: https://arxiv.org/abs/2605.12096 (дата обращения: 06.10.2026).
24. Онаркулов М.К., Юсупов М.А., Умиржонов Л.А. Наборы данных жестовых языков мира: комплексный обзор от западных до центральноазиатских языков // Научно-технический журнал ФерГТУ. 2026. Т. 30, спец. вып. № 1. С. 195–199.
25. Tricco A.C., Lillie E., Zarin W. [et al.] PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation // Annals of Internal Medicine. 2018. Vol.

(cc) BY

169, No. 7. P. 467–473. DOI: 10.7326/M18-0850.
26. Page M.J., McKenzie J.E., Bossuyt P.M. [et al.] The PRISMA 2020 statement: an updated guideline for reporting systematic reviews // BMJ. 2021. Vol. 372. Art. n71. DOI: 10.1136/bmj.n71.
27. Kvanchiani K., Surovtsev P., Nagaev A. [et al.] Bukva: Russian Sign Language Alphabet [Электронный ресурс] // arXiv. 2024. URL: https://arxiv.org/abs/2410.08675 (дата обращения: 06.10.2026).
28. Imashev A., Mukushev M., Kimmelman V., Sandygulova A. A Dataset for Linguistic Understanding, Visual Evaluation, and Recognition of Sign Languages: The K-RSL // Proceedings of the 24th Conference on Computational Natural Language Learning (CoNLL 2020). Online: ACL, 2020. P. 631–640. DOI: 10.18653/v1/2020.conll-1.51.
29. Mukushev M., Ubingazhibov A., Kydyrbekova A. [et al.] FluentSigners-50: A signer independent benchmark dataset for sign language processing // PLOS ONE. 2022. Vol. 17, No. 9. Art. e0273649. DOI: 10.1371/journal.pone.0273649.
30. Alishzade N., Hasanov J. AzSLD: Azerbaijani Sign Language Dataset for Fingerspelling, Word, and Sentence Translation with Baseline Software // Data in Brief. 2025. Vol. 58. Art. 111230. DOI: 10.1016/j.dib.2024.111230.
31. O'Brien C., Sant G., Müller M., Ebling S. Evaluation of Pose Estimation Systems for Sign Language Translation // Proceedings of the LREC 2026 12th Workshop on the Representation and Processing of Sign Languages: Language in Motion. Palma, Mallorca: ELRA, 2026. P. 371–386. URL: https://aclanthology.org/2026.signlang-1.39/ (дата обращения: 06.10.2026).
32. Selvaraj P., NC G., Kumar P., Khapra M. OpenHands: Making Sign Language Recognition Accessible with Pose-based Pretrained Models across Languages // Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. Vol. 1: Long Papers. Dublin: ACL, 2022. P. 2114–2133. DOI: 10.18653/v1/2022.acl-long.150.
33. Bilge Y.C., Ikizler-Cinbis N., Cinbis R.G. Cross-lingual few-shot sign language recognition // Pattern Recognition. 2024. Vol. 151. Art. 110374. DOI: 10.1016/j.patcog.2024.110374.
34. Kindiroglu A.A., Kara O., Ozdemir O., Akarun L. Transfer Learning for Cross-dataset Isolated Sign Language Recognition in Under-Resourced Datasets // 2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition (FG 2024). IEEE, 2024. DOI: 10.1109/FG59268.2024.10582029.
35. Ryumin D., Ivanko D., Axyonov A. Cross-language transfer learning using visual information for automatic sign gesture recognition // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2023. Vol. XLVIII-2/W3-2023. P. 209–216. DOI: 10.5194/isprs-archives-XLVIII-2-W3-2023-209-2023.
36. Peng K., Yin C., Zheng J. [et al.] Navigating Open Set Scenarios for Skeleton-based Action Recognition // Proceedings of the AAAI Conference on Artificial Intelligence. 2024. Vol. 38, No. 5. DOI: 10.1609/aaai.v38i5.28247.
37. Yang L., Yan W., He Y. [et al.] SAME: Signer-Aware Mixture-of-Experts for Test-Time Adaptation in Sign Language Translation // Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics. Vol. 1: Long Papers. 2026. P. 21274–21289. DOI: 10.18653/v1/2026.acl-long.973.
38. Li Y., Ren B., Hu K. [et al.] KD-MSLRT: Lightweight Sign Language Recognition Model Based on Mediapipe and 3D to 1D Knowledge Distillation // Proceedings of the AAAI Conference on Artificial Intelligence. 2025. Vol. 39, No. 27. P. 28177–28185. DOI: 10.1609/aaai.v39i27.35037.
39. Javorek V., Honzík J., Gruber I. [et al.] Sign Language Recognition in the Age of LLMs // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2026). 2026. P. 10511–10519.
40. Рюмин Д.А. Модели и методы автоматического распознавания элементов русского жестового языка для

(cc) BY *dtai.uz*

человеко-машинного взаимодействия: дис. … канд. техн. наук: 05.13.17. СПб.: Университет ИТМО, 2020.
41. Муханов С.Б. Разработка и применение высокоточных методов распознавания образов: дис. … д-ра философии (PhD): 6D070400. Алматы: Международный университет информационных технологий, 2024.
42. Zhassuzak M., Buribayev Z., Aouani M. [et al.] Bidirectional Kazakh Sign Language prosody-aware translation using computer vision and speech recognition techniques // Frontiers in Artificial Intelligence. 2026. Vol. 9. Art. 1835419. DOI: 10.3389/frai.2026.1835419.
43. Бюллетень Высшей аттестационной комиссии при Академии наук Республики Узбекистан. 2026. № 1 [Электронный ресурс]. URL: https://oak.uz/userfiles/files/2026-1-son.pdf (дата обращения: 06.10.2026).
44. Kayumov O.A., Kayumova N.R. Based on Machine Learning Algorithms to Recognize Uzbek Sign Language (UzSL) [Электронный ресурс] // Zenodo. 2023. DOI: 10.5281/zenodo.10093343.
45. Kayumov O.A., Kayumova N.R. Algorithm and model of Latin alphabet recognition in Uzbek Sign Language (UzSL) based on angular signs obtained from finger joint and wrist coordinates // International Journal of Artificial Intelligence. 2024. Vol. 4, No. 3. P. 628–633. URL: https://www.academicpublishers.org/journals/index.php/ijai/article/view/745 (дата обращения: 06.10.2026).
46. Kayumov O.A., Kayumova N.R., Xodjabekova D.F. Uzbek Sign Language Classifier Based on Machine Learning // European International Journal of Multidisciplinary Research and Management Studies. 2024. Vol. 4, No. 5. DOI: 10.55640/eijmrms-04-05-43.
47. Kayumov O.A., Kayumova N.R., Xodjabekova D.F. Development of a Sign Language Recognition Model for Uzbek Words Using Deep Learning Methods // International Multidisciplinary Journal for Research & Development. 2024. Vol. 11, No. 5. URL: https://www.ijmrd.in/index.php/imjrd/article/view/1568 (дата обращения: 06.10.2026).
48. Akhmedov F., Umarov M., Urinov E., Turgunov B. An Algorithm of Recognition Uzbek Sign Language // 2024 International Conference on Information Science and Communications Technologies (ICISCT). IEEE, 2024. DOI: 10.1109/ICISCT64202.2024.10956783.
49. Urinov E.M., Umarov M.A. O‘zbek ishora tili harflarini tanib olish algoritmi // Al-Farg‘oniy avlodlari. 2024. № 1(5). B. 29–33. DOI: 10.5281/zenodo.10867054.
50. Jovliyeva S. A Webcam-Based Recognition System for the Uzbek Fingerspelling Alphabet // Innovative: International Multidisciplinary Journal of Applied Technology. 2026. Vol. 4, No. 3. DOI: 10.51699/x67ey438.
51. Saroliya A., Sabitjanov M., Avazov X. Assistance to Uzbek people with speaking disabilities using Machine Learning and Sign Language Detection // 2024 IEEE 9th International Conference for Convergence in Technology (I2CT). 2024. P. 1–5. DOI: 10.1109/I2CT61223.2024.10544372.
52. Sabitjanov M., Avazov X., Isomitdinova S., Srivastava A.K. A Machine Learning Algorithm to Assist People with Speaking Disabilities for Sign Language Detection of People Speaking Uzbek Language // 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS). Tashkent, 2024. P. 768–772. DOI: 10.1109/ICTACS62700.2024.10841300.
53. Musaev M., Juraev D., Ochilov M., Abdullaeva M. Text processing technology in Uzbek speech to sign language translation systems // AIP Conference Proceedings. 2024. Vol. 3244. Art. 030062. DOI: 10.1063/5.0241738.
54. Juraev D.B., Ochilov M.M., Abdullaeva M.I. Semantically Enhanced Multimodal Neural Architecture for Uzbek Sign Language Translation // Advanced Computing: An International Journal. 2025. Vol. 16, No. 6. P. 9–27. DOI: 10.5121/acij.2025.16602.
55. Ergashev J., Rakhmatova M., Eshimov M. SIGNLY.UZ: A Gamified Mobile Platform for Uzbek Sign Language Learning and

(cc) BY *dtai.uz*

Offline Recognition [Электронный ресурс] // Zenodo. 2026. DOI: 10.5281/zenodo.22814776.

56. Ergashev J., Rakhmatova M., Eshimov M. Comparative Evaluation of CNN-LSTM and Squeezeformer-Transformer Architectures for Real-Time Uzbek Sign Language Recognition // Innovation Science and Technology. 2026. Vol. 4, No. 9. P. 307–315. DOI: 10.5281/zenodo.22840569.
57. Henkel C. Google — American Sign Language Fingerspelling Recognition: 1st place solution [Электронный ресурс] // GitHub. 2023. URL: https://github.com/ChristofHenkel/kaggle-asl-fingerspelling-1st-place-solution (дата обращения: 06.10.2026).
58. UZSL Mobile Translator — O‘zbek Imo-Ishora Tili Tarjimoni [Электронный ресурс]: программный репозиторий // GitHub. 2026. URL: https://github.com/samuraix9179/surdo_Uzsl (дата обращения: 06.10.2026).
59. Hendrycks D., Gimpel K. A Baseline for Detecting Misclassified and Out-of-Distribution Examples in Neural Networks [Электронный ресурс] // 5th International Conference on Learning Representations (ICLR 2017). 2017. URL: https://arxiv.org/abs/1610.02136 (дата обращения: 07.10.2026).

(cc) BY *dtai.uz*