

TOWARDS BETTER RUSSIAN SIGN LANGUAGE TRANSLATION: DATASET CHALLENGES AND TRAINING STRATEGIES

Z. V. Rybinskaia

The article is presented by the scientific supervisor – Candidate of Technical Sciences, Associate Professor Khlopotov M. V.

This paper analyzes the current state of Russian Sign Language (RSL) datasets and identifies the main challenges in their use for AI-based translation systems. A comparative analysis of the Bukva, Slovo, and HaGRID datasets was conducted to assess their size, annotation quality, diversity, and openness. The results show that while Slovo provides the most linguistically diverse material, Bukva and HaGRID contribute valuable methodological and validation frameworks. The study emphasizes the need for hybrid and multimodal datasets to enhance inclusivity and practical RSL translation performance.

Keywords: Russian Sign Language, dataset analysis, artificial intelligence, gesture recognition, machine translation.

In recent years, artificial intelligence technologies have been increasingly applied to promote inclusive communication. Examples include sign language recognition and translation, automatic subtitling, and interactive assistants for people with hearing impairments. When developing a Russian-language intelligent service aimed at overcoming communication barriers, one of the key tasks is the creation and adaptation of translation models for Russian Sign Language (RSL). This study aims to identify the key challenges in existing RSL-related datasets and propose practical strategies for improving their quality and applicability in translation models. Unlike most previous studies that focus on English-based systems, this paper highlights the underexplored Russian context, where data accessibility and linguistic specificity remain key limitations.

To evaluate the strengths and weaknesses of existing datasets, a comparative analysis was conducted across three major Russian gesture datasets, Bukva, Slovo, and HaGRID, based on four key criteria:

- 1) dataset size and coverage,
- 2) diversity of signers and recording conditions,
- 3) annotation and validation quality,
- 4) openness of licensing.

To complement the dataset analysis, previous research on Russian dactyl recognition using neural networks, particularly YOLO-based architectures, was also considered [1-3]. This helps illustrate how object detection methods can support gesture segmentation and labeling in RSL datasets.

One of the most notable resources is the Bukva dataset, created by the AI Forever team. It contains approximately 27 GB of video data (3,757 RGB videos) with both dynamic and static representations of the RSL manual alphabet – each letter is represented by more than 100 examples, divided into training and test subsets. The open-source nature of Bukva allows its use for further fine-tuning of neural networks, as well as for correction and expansion of the dataset. However, despite its integrity and accessibility, the dataset remains limited in size and semantic depth: it contains only alphabetic gestures, without words, phrases, or consideration of communicative context and signer variability. This restricts the applicability of fine-tuned models in real-life communication scenarios.

The next step in development is represented by the Slovo dataset, released by SberDevices. It includes 20,400 Full HD videos, over 1,000 unique gesture classes

(including dactyl alphabet signs), recorded by 194 native signers and RSL experts from around the world. This resource represents a significant breakthrough: a large number of gesture classes, diverse signers, and high recording quality all contribute to better model training conditions. Special attention was paid to data accuracy – the complete dataset creation cycle included data collection, validation, annotation, and even a testing phase to identify users proficient in RSL. Nevertheless, Slovo also faces certain limitations: recordings were made in controlled environments, many gestures are static, and the dataset does not yet cover full sentences or natural discourse used in spontaneous communication. Moreover, only part of the dataset and several fine-tuned neural models were released publicly. However, such limitations are difficult to avoid when forming a conventional dataset of this type.

Another noteworthy resource is HaGRID, another large-scale gesture dataset developed by SberDevices, though focused not on RSL translation but on gesture control tasks. It contains 554,800 images across 18 gesture classes collected from 37,583 users worldwide. HaGRID deserves attention due to its rigorous multi-stage validation process, which makes it one of the cleanest and most structured datasets in Russian practice. To minimize errors and eliminate ambiguous data, the developers divided validation into two stages. The first stage included baseline filtering and annotator training, while the second involved re-checking images under stricter constraints: removing photos of children, unclothed individuals, and any images containing textual inscriptions or symbols (national, religious, etc.). The overlap between validation sets was strictly limited to five samples, as the cost of error was considered particularly high. Although HaGRID does not cover the linguistic specifics of RSL (phonology, syntax, or cultural variation), it remains a valuable general resource for gesture recognition. Its thorough validation and data col-

lection pipeline demonstrate the maturity of Russian data-engineering approaches and may serve as a methodological foundation for the creation of specialized RSL datasets in the future.

Thus, the Russian dataset landscape for RSL is still in an active stage of formation, striving toward completeness and quality. The main challenges can be summarized as follows:

- 1) limited dataset volume and gesture class diversity, especially for phrasal or connected sign language speech;
- 2) insufficient variability in signers, backgrounds, and recording conditions, which reduces model generalization;
- 3) complex annotation requirements due to multimodality (hand movement, body posture, facial expression, and context), which are not yet fully addressed;
- 4) underdeveloped standards for openness, licensing, and interoperability;
- 5) significant financial and time costs of collecting video data from RSL users and experts.

The analysis shows that while Slovo provides the most diverse and scalable data, Bukva remains limited to alphabetic signs, and HaGRID, though methodologically robust, lacks linguistic specificity for RSL translation. Overall, Russian datasets for sign language research are still in a formative stage and require high-quality and representative datasets, yet the Russian implementation of this task still faces serious challenges – limited data volumes, low variability, annotation complexity, and high production costs. These findings suggest the need for a hybrid dataset combining the linguistic richness of Slovo with the methodological rigor of HaGRID, along with extended annotation frameworks covering facial expressions, body posture, and communicative context. Future research may focus on developing multimodal RSL datasets that better reflect real-life communication and promote inclusive, AI-driven translation services in the Russian context.

References

1. Dolin. A. A. Neural network concept for recognizing elements of the Russian dactyl alphabet // International Journal Symbol of Science. 2024. №. 2. pp. 42-47.
2. Kharlamenkov A. E., Ashrafi A., Mokhnachyov V. S., Filippovich U. N., Chernenko S. E. Sign language recognition based on kinematic segmentation of gestures // Contemporary Pedagogical Education. 2022. №. 2. pp. 126-131.
3. Pechenov I. P., Gulyaev I. V., Shabanova I. N., Astrelin D. M. Development of models and methods for recognizing the Russian dactyl alphabet [Electronic resource] // Engineering Bulletin of Don. 2024. №. 5 (19.03.2024). URL: <http://www.ivdon.ru/ru/magazine/archive/n5y2024/9212> (last request: 05.11.2025).

Rybinskaia Zlata Vladislavovna – student, St. Petersburg State University of Information Technology, Mechanics and Optics, (St. Petersburg, Russia), zltrbn@gmail.com

Khlopotov Maxim Valerievich – scientific supervisor – Candidate of Technical Sciences, Associate Professor, Faculty of Applied Informatics, St. Petersburg State University of Information Technology, Mechanics and Optics, (St. Petersburg, Russia), khlopotov@itmo.ru

К ВОПРОСУ О СОЗДАНИИ УЛУЧШЕННОГО ПЕРЕВОДА РУССКОГО ЖЕСТОВОГО ЯЗЫКА: ПРОБЛЕМЫ НАБОРОВ ДАННЫХ И СТРАТЕГИИ ОБУЧЕНИЯ

З. В. Рыбинская

Статья представлена научным руководителем – к. т. н., доц. Хлопотовым М. В.

*В данной статье анализируется текущее состояние наборов данных по русскому жестовому языку (РЖЯ) и выявляются основные проблемы их использования в системах перевода на основе искусственного интеллекта. Проведён сравнительный анализ наборов данных *Vukva*, *Slovo* и *HaGRID* для оценки их размера, качества аннотации, разнообразия и открытости. Результаты показывают, что, хотя *Slovo* предоставляет наиболее лингвистически разнообразный материал, *Vukva* и *HaGRID* вносят ценный вклад в методологическую основу и валидационные механизмы. Исследование подчёркивает необходимость гибридных и мультимодальных наборов данных для повышения инклюзивности и практической эффективности перевода РЖЯ.*

Ключевые слова: русский жестовый язык, анализ наборов данных, искусственный интеллект, распознавание жестов, машинный перевод.

Список источников и литературы

1. Долин А. А. Нейросетевая концепция распознавания элементов русской дактильной азбуки // Символ науки: международный научный журнал. 2024. №. 2. С. 42-47.
2. Харламенков А. Е., Ашрафи А., Мохначев В. С., Филиппович У. Н., Черненко С. Е. Распознавание жестового языка на основе киремного деления жестов // Современное педагогическое образование 2022. №. 2. С. 126-131.
3. Печенов И. П., Гуляев И. В., Шабанова И. Н., Астрелин А. А., Курляк Д. М. Разработка моделей и методов распознавания русской дактильной азбуки // Инженерный вестник Дона 2024. №. 5 (19.03.2024). URL: <http://www.ivdon.ru/ru/magazine/archive/n5y2024/9212> (дата обращения: 05.11.2025).

Рыбинская Злата Владиславовна – студентка, Санкт-Петербургский государственный университет информационных технологий, механики и оптики, (г. Санкт-Петербург, Россия), zltrbn@gmail.com

Хлопотов Максим Валерьевич – научный руководитель – к. т. н., доцент, факультет прикладной информатики, Санкт-Петербургский государственный университет информационных технологий, механики и оптики, (г. Санкт-Петербург, Россия), khlopotov@itmo.ru

Статья поступила в редакцию: 12.05.2026; принята к публикации: 30.05.2026

ДЛЯ ЦИТИРОВАНИЯ:

Рыбинская З. В. К вопросу о создании улучшенного перевода русского жестового языка: проблемы наборов данных и стратегии обучения // Социогуманитарные коммуникации. – 2026. – № 2(16). – С. 203-206.

FOR CITATION:

Rybinskaia Z. V. K voprosu o sozdanii uluchshennogo perevoda russkogo zhestovogo yazyka: problemy naborov dannyh i strategii obucheniya [Towards better Russian sign language translation: dataset challenges and training strategies] // Sociogumanitarnye kommunikacii [Social and humanitarian communications]. 2026. No 2(16). P. 203-206.