Creating the First Bengali-Russian Sign Language Dictionary for Inclusive Multilingual Communication

Arifa Ashrafi, Yuriy Philippovich

Abstract — **This paper outlines the creation of the first Bengali-Russian Sign Language Dictionary, aimed at promoting inclusive communication by a multilingual, user-friendly platform. The research involved creating a detailed dataset with static and dynamic samples for both Bengali and Russian sign languages, supported by the implementation of different Convolutional Neural Network (CNN) models for effective sign recognition. A user interface for Bengali-Russian Sign Languages has been developed to offer an accessible experience, with expert evaluations and performance efficiency confirming its usability and user-friendliness. The dictionary's design supports future scalability, allowing for additional languages and signs to be added. By integrating advanced CNN models, the platform achieves robust accuracy, making it suitable for various educational and community settings. This work lays a foundation for further innovations in digital inclusivity, potentially transforming sign language education for hearing impaired communities especially for underprivileged individuals. This research also explores video-to-animation conversion techniques, resulting in improved feedback for users during the learning process. To understand the educational impact on diverse student groups, field studies were conducted with both street and mainstream children. The results showed that a positive, digitally inclusive learning environment greatly improves engagement and understanding. Also the findings from the research suggest that once the dictionary is launched, this platform will significantly bridge communication gaps and provide a structured educational resource for Bengali, and Russian hearing-impaired communities.**

Keywords — **Bengali Sign Language, Russian Sign Language, Sign Language recognition, video based sign language dictionary, machine learning.**

I. Introduction

Communication is fundamental to human communication. For hearing-impaired communities, language barriers often create obstacles to full societal integration. Sign languages are the most natural and meaningful way of communication for hearing impaired people [1], yet there remains a lack of accessible, multilingual resources that support sign language learning and translation across diverse languages. In particular, Bengali and Russian sign languages are underrepresented in digital learning platforms, reducing the opportunities for people who use these languages to fully participate in educational and social activities. In order to facilitate the learning of sign languages by the hearing impaired community, researchers have compiled sign language repositories consisting of gestures [2]. This research addresses this gap by developing the first Bengali-Russian Sign Language Dictionary, a video-based, user-centered platform designed to support accessible, inclusive sign language education.

To achieve this, an extensive video and image dataset of Bengali and Russian sign languages, enabling robust training and testing for sign language recognition models have been created. For dynamic signs, different video classification methods such as Dynamic Time Warping (DTW), Hidden Markov Model (HMM), and CNN-LSTM Architecture have been modified, and used [3, 4, 5]. Through implementing advanced Convolutional Neural Networks (CNNs) and MediaPipe technology, the platform achieves high recognition accuracy and responsiveness in sign language detection and translation. Additionally, by converting videos into animations, the platform offers clearer, more engaging feedback to users, enhancing the learning experience.

The research extends beyond technical implementation, incorporating a study of educational environments to assess how learning conditions impact engagement and effectiveness. Field studies with street children and mainstream students provided insights into how competitive, technologically advanced settings can foster greater motivation and skill development in children with diverse needs. Findings indicate that an inclusive educational approach, supported by user-friendly technology, can significantly enhance learning outcomes for both hearing and hearing-impaired students.

This paper discusses the development of the Bengali-Russian Sign Language Dictionary, focusing on the methodology, design, and evaluation processes. Through expert evaluations, we refined the user interface to ensure accessibility and ease of use, particularly for those with hearing impairments. Our platform is poised to launch as a pioneering educational tool, contributing to the global effort to bridge communication divides and promote equitable access to language resources for all.

II. Creating the First Bengali-Russian Sign Language Dictionary

A. Advancing Multilingual Accessibility

The creation of the first Bengali-Russian Sign Language

Manuscript received January 23, 2025.
This work was supported in part by Moscow Polytechnic University under V.E. Fortov's Grant.

A. Ashrafi – Moscow Polytechnic University (email: arifaa13@ gmail.com)
Y.N. Philippovich – Moscow Polytechnic University (email: y_philippovich@mail.ru)

Dictionary marks a significant advancement in multilingual accessibility for the hearing-impaired community. This resource not only bridges communication gaps but also fosters cultural and linguistic exchange between Bengali and Russian speakers. By addressing the unique challenges of sign language translation, this dictionary paves the way for more inclusive, cross-cultural interactions and serves as a foundation for future language accessibility tools.

B. Problem Statement

Despite the significant advancements in technology, accessibility to sign language learning tools remains limited, especially in developing countries like Bangladesh. Existing tools for sign language education are often fragmented, and lack the interactivity that needs to be addressed for the target community. The gap in resources, particularly for Bengali sign language, highlights the need for an inclusive and comprehensive platform that bridges this communication gap.

III. RESEARCH METHODOLOGY

This research adopts an experimental approach to the design, development, and evaluation of a sign language dictionary for Bengali and Russian sign languages. The methodology is defined in figure 1.

The methodology for developing the Bengali-Russian Sign Language Dictionary begins with requirements analysis to identify essential features, using feedback from hearing-impaired users, educators, and sign language experts. This feedback informs the platform's design, ensuring a multilingual, user-friendly, and accessible interface.

In data collection, an extensive video and image dataset of Bengali, and Russian, signs are compiled with expert collaboration. Signs are annotated with metadata to support machine learning model development using CNNs and MediaPipe for real-time hand and body tracking, optimized through data augmentation for accurate recognition across languages.

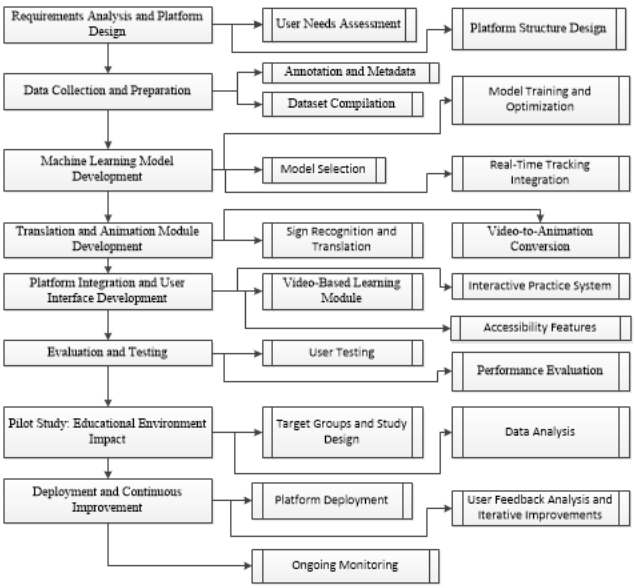

Fig.1 Process of motivating students' to participate in the international conference designed to improve their adaptation ability for global trends

Translation and animation modules convert recognized signs into translated text and animations, creating an engaging learning experience. The platform integration phase combines the model with interactive features for video-based learning and practice, enhanced by accessibility options like captions and keyboard navigation.

Evaluation and testing involve usability and performance assessments, ensuring accuracy, reliability, and user satisfaction. The pilot study observes the platform's impact on various educational groups, showing that inclusive technology supports improved learning outcomes.

Finally, deployment and continuous improvement maintain platform performance, with regular updates based on user feedback to enhance recognition accuracy, expand the dataset, and meet the evolving needs of a diverse user base.

IV. STAGES FOR DICTIONARY DEVELOPMENT

The development of the Bengali-Russian Sign Language Dictionary involves a structured, multi-stage workflow, as shown in the flowchart in figure 2.

The project's front end is created using HTML, CSS, and JavaScript to ensure a convenient and user-friendly interface.

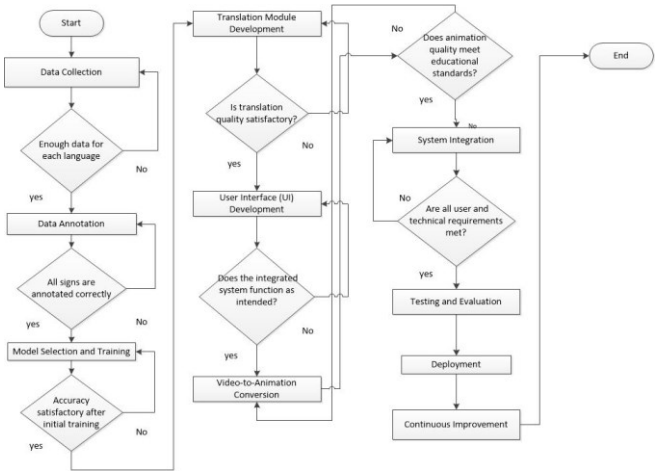

Fig.2 Flowchart Stages for Dictionary Development

The backend uses Node.js to manage data processing and to handle the dictionary's complex sign language datasets, facilitating communication between the front-end interface and the database.

V. DATASET DEVELOPMENT

A. Bengali Sign Language

The dataset development for the Bengali Sign Language Dictionary was a crucial aspect of this research. It involved the collection of diverse video (sample shown in figure 3) and image data representing signs from Bengali sign languages, with the goal of capturing a wide range of hand movements, facial expressions, and contextual nuances.

Fig.3 Random samples of video datasets of Bengali Sign Language

B. *Russian Sign Language*

The website of the All-Russian Union of Russian Sign Language Translators "SurdoCenter" (http://surdocentr.ru) contains the Electronic reference and analytical system of Russian Sign Language. There are over 2000 original videos of signs available for researchers, scholars, hearing-impaired, and so on [6]. These videos (random sample shown in figure 4) have been used for Russian Sign Language dataset development.

To support the creation of a detailed dataset many researchers commonly utilize skin-color detections [7–9] and haar-like features [10] for hand region segmentation. To more accurately find and extract the hand region, Kinect depth cameras can be used [11, 12, 13, 14, 15]. But these approaches necessitate additional sensor costs, which is one of the key reasons why scientists and researchers have focused on AI-related approaches for sign language recognition [3].

For this research work the dataset development was done in collaboration with sign language experts to ensure linguistic and cultural accuracy, providing a solid foundation for the recognition and translation features of the dictionary.

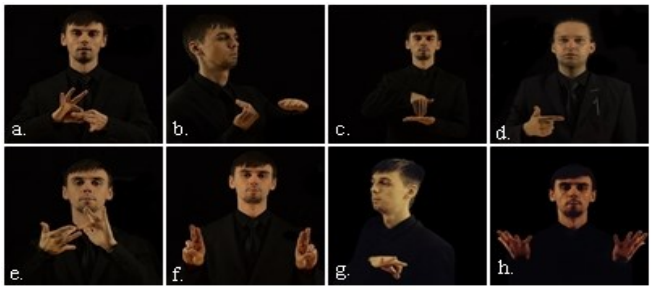

Fig.4 Random samples of video datasets of Russian Sign Language: a - верный (faithful), b - вызов (call), c - город (city), d - грамотный (literate), e - интернет (internet), f - правило (rule), g - прибыль (profit), h - светлый (light) [3].

VI. IMPLEMENTATION

The implementation of the Bengali-Russian Sign Language Dictionary involved several key stages. First, video and image datasets for both languages were collected and annotated with relevant metadata, including categories like alphabet, country names, and commonly used phrases, to structure the learning process. Three user interfaces were developed:

1) Text Input Interface: Users manually type text, which the dictionary then translates to the target sign language.
2) Selectable Interface: Users choose from a predefined list of common words or phrases.
3) Categorized Interface: A structured interface allowing users to browse signs by categories such as days, months, or alphabet.

A real-time sign language recognition system was built using CNNs and integrated with MediaPipe for hand tracking, though limited resources for Bengali sign language impacted accuracy for certain gestures. Expert evaluations and user feedback led to refinements, particularly in the categorized UI. A video-to-animation module was implemented, enhancing the learning experience by providing clearer visual feedback and allowing users to more accurately mimic signs.

This research methodology follows a systematic approach to convert sign language videos into animations using video tracking technology. It is divided into several key steps to ensure the effective development and testing of a multilingual sign language dictionary. The focus is on Bengali, Russian, and American Sign Languages.

For video tracking and motion capture several tools are available, like MediaPipe [17, 18, 19, 20], OpenPose, or similar body-tracking technology. The reason to choose MediaPipe technology is its wide use in applications that require precise hand (shown in Figure 2), face, and body tracking. Studies [4] highlight its efficiency in detecting complex gestures in real time due to its high precision and performance across multiple devices, from mobile phones to desktops. The framework is especially effective for tasks like sign language recognition, where accurate tracking of finger and hand motions is vital for proper interpretation [21].

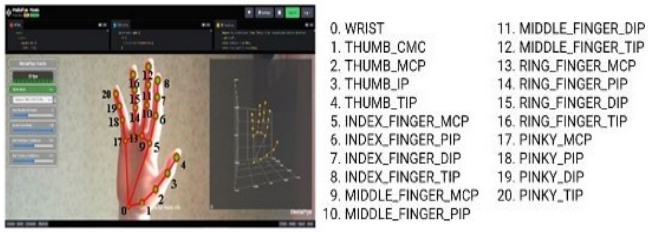

Fig.5 MediaPipe Hand Tracking [6]

The process of utilizing MediaPipe to convert sign language videos into animations involves the proposed steps defined in Figure 6.

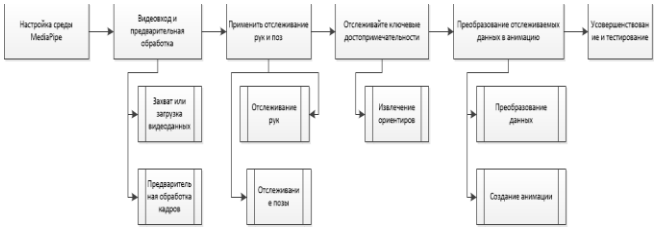

Fig.6 The process of utilizing MediaPipe to convert sign language videos into animations [23].

The sign language dictionary mockup shown in figure 7 allows users to choose between two language options: Russian and Bengali, catering to diverse language needs. At the center of the interface, a cartoon of children adds an engaging and friendly touch, designed to make the tool appealing, especially to young users.

Fig.7 Mockup of proposed Bengali-Russian Sign Language dictionary (Web Application).

To improve the translator shown in figure 8, the size of the datasets will be increased. The proposed algorithm can be updated in the future to meet the user requirements and market analysis [24].

The implementation of multilingual sign language text translation dictionaries shows a very positive result when analyzed by more than fifty users [25].

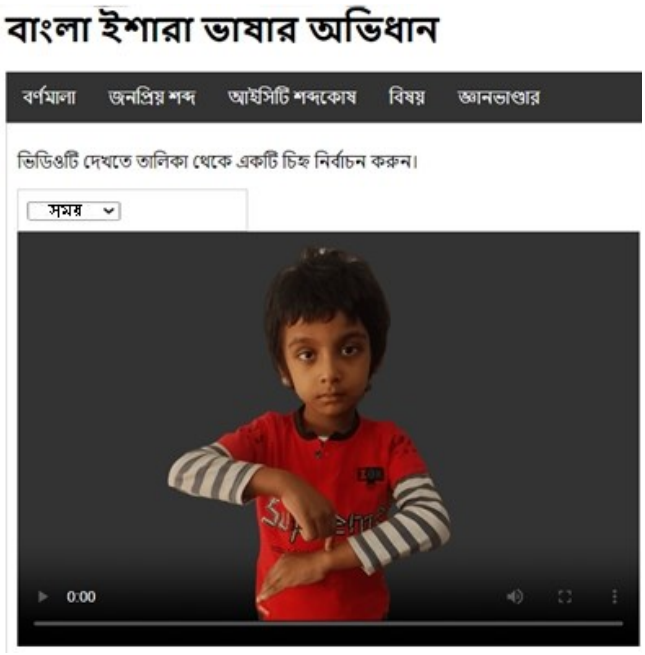

Fig.8 Bengali-Russian Sign Language Dictionary Interface – Bengali Section

To evaluate the usability of the multilingual sign language translator, the factorial expert evaluation method (Figure 9), also known as heuristic evaluation, is used. This method has less complexity than others.

Fig.9 The process of conducting factorial expert evaluation

All experts will give scores for each criterion in the range from 3 to 5. Here, j = 1 ... n, representing experts, and i = 1 ... n, representing factors [21].

TABLE I. Displaying the relation [i, j]

i/j	j_1	j_2	j_3	…	j_n
i_1	$[i_1, j_1]$	$[i_1, j_2]$	$[i_1, j_3]$	…	$[i_1, j_n]$
i_2	$[i_2, j_1]$	$[i_2, j_2]$	$[i_2, j_3]$	…	$[i_2, j_n]$
i_3	$[i_3, j_1]$	$[i_3, j_2]$	$[i_3, j_3]$	…	$[i_3, j_n]$
…	…	…	…	…	…
i_n	$[i_n, j_1]$	$[i_n, j_2]$	$[i_n, j_3]$	…	$[i_n, j_n]$

The sum of expert points/assessments is determined by the formula:

$$S_j = \sum_{i=n}^{i=1} \quad f[i, j] \tag{1}$$

The web application of text to sign language translator shown in Figure 8 is selected as the most user-friendly for Bengali Sign Language after using the expert evaluation method. In addition, experts provided their analysis to improve the usability, which will be used to further improve the web application.

VII. Results and Analysis

System Accuracy: For understanding system accuracy suppose, S_B and S_R represent the recognition accuracy for Bengali and Russian sign languages, respectively. The accuracy for each sign language (S_L) can be calculated as:

$$S_L = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} * 100 \tag{2}$$

For Bengali sign language S_B:

$$S_B = \frac{C_B}{T_B} * 100 \tag{3}$$

Where, C_B = Number of correctly recognized signs for Bengali, T_B = Total number of test signs for Bengali.

For Russian sign language S_R:

$$S_R = \frac{C_R}{T_R} * 100 \tag{4}$$

Where, C_R = Number of correctly recognized signs for Russian, T_R = Total number of test signs for Russian.

Here Russian Sign Language shows better results than Bengali Sign Language.

User Interface Performance: Suppose, UI_{text}, UI_{select}, and $UI_{categorize}$ represent the user satisfaction scores for the text input, selectable, and categorized user interfaces, respectively. User satisfaction can be quantified based on expert evaluations and user surveys. And let $U_{feedback}$ represent the satisfaction score, which ranges from 0 to 1, where 1 indicates the highest satisfaction.

$$U_{feedback} = \frac{\sum_{i=1}^{n} \quad user\ satisfaction\ score}{n} \tag{5}$$

Where, n = Total number of users participating in feedback, user satisfaction score = Rating for each user, ranging from 0 to 1.

The categorized interface had higher ratings = 0.85, reflects the satisfaction with the categorized interface based on user feedback.

The results of the study indicate the effectiveness of the proposed sign language dictionary. The categorization-based interface, which grouped words into meaningful sections, received the best user feedback for Bengali sign language.

Text input-based interfaces were less preferred due to the complexity of Bengali sign language gestures.

Real-Time Recognition Performance: The performance of the model is evaluated based on accuracy A, precision P , recall R , and $F1$ score:

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$P = \frac{TP}{TP+FP} \quad (7)$$

$$R = \frac{TP}{TP+FN} \quad (8)$$

$$F1 = 2 * \frac{P * R}{P + R} \quad (9)$$

where TP, TN, FP, and FN [16, 22] represent true positives, true negatives, false positives, and false negatives, respectively. High F1 scores across all three languages indicate effective cross-language recognition and translation.

1000 gestures for each language using MediaPipe for motion tracking. The accuracy, precision, and F1 score calculated for Bengali and Russian sign languages as follows:

TABLE II. Experimental results for Real-Time Recognition Performance of bengali (BSL), and russian (RSL) sign languages with average of one hundred epochs

Real-Time Recognition Performance	Bengali Sign Language	Russian sign language
Accuracy	0.754	0.953
Precision	0.931	0.974
recall	0.793	0.983
f1 score	0.854	0.973

The real-time recognition system using MediaPipe achieved a 75% accuracy for Bengali sign language and 95% accuracy for Russian sign language.

To explain user involvement and make a comparison between video-based and animation-based sign language dictionaries, metrics like engagement time, Comprehensive Accuracy, interaction frequency, and Learning Retention have been used. The experiment and analysis of Interaction frequency have been described below:

Interaction frequency refers to the number of interactions (e.g., pausing, replaying, slowing down) a user makes with the dictionary while learning. Let I_v and I_a represent the interaction frequencies for the video-based and animation-based systems.

If F_v and F_a are the total number of interactions in the video-based and animation-based dictionaries, and N_v and N_a are the total number of users, the average interaction frequency per user can be calculated as:

$I_v = \frac{F_v}{N_v}$(Video-based interaction frequency per user) (10)

$I_a = \frac{F_a}{N_a}$(Animation-based interaction frequency per user) (11)

Higher interaction frequency in the animation-based system, I_a compared toI_v, indicate that users interact more with animations, likely due to greater control over the learning process [19].

VIII. FUTURE PROGRESSION

For future work, the following steps are proposed:

1) Expansion of Language Support: Extending the dictionary to include additional sign languages to cater to a wider audience.
2) Improvement in Gesture Recognition: Enhancing the real-time recognition accuracy, especially for Bengali sign language, through the use of more diverse datasets and improved CNN models.
3) Interactive Learning Tools: Developing additional interactive learning features, such as quizzes and games, to make the learning process more engaging.
4) Inclusive Education Integration: Exploring how the dictionary can be integrated into inclusive educational settings, offering children with hearing impairments more personalized and interactive learning experiences.

IX. Conclusion

This research introduces the first Bengali-Russian Sign Language Dictionary, a tool that supports inclusive, multilingual education through sign language. The dictionary uses a video-based platform, and is supposed to be used with advanced technologies like CNN and MediaPipe, which help to achieve accurate sign recognition and create an interactive experience for users. It is designed to be user-friendly, making it suitable for children, hearing-impaired individuals, and their peers.

Our study shows that digital tools like this dictionary can positively impact learning and engagement, especially for less-represented languages such as Bengali. By providing an accessible, categorized interface, this platform helps bridge communication gaps for the Bengali and Russian hearing-impaired communities.

Future work will focus on expanding the dictionary to support additional languages and improving recognition accuracy, so it becomes even more effective as an educational tool. This project demonstrates the importance of multilingual resources for promoting cross-cultural understanding and accessibility. It serves as a foundation for future innovations aimed at improving the hearing-impaired community with technology.

Acknowledgements

We express our gratitude to the All-Russian Union of Russian Sign Language Translators, "SurdoCenter," for providing invaluable resources and guidance on Russian sign language. We also extend special thanks to our collaborators and expert advisors in Bengali sign language who offered critical insights and feedback throughout the project.

We are deeply appreciative of the support from our educational institutions, as well as the children, educators, and volunteers who participated in field studies. Their engagement and feedback have been instrumental in refining the platform.

References

[1] Ashrafi, A., Mokhnachev, V.S., Philippovich, Y.N., Tsilenko, L.P. (2020). Development of Image Dataset Using Hand Gesture Recognition System for Progression of Sign Language Translator. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds) Software Engineering Perspectives in Intelligent Systems. CoMeSySo 2020. Advances in Intelligent Systems and Computing, vol 1294. Springer, Cham. https://doi.org/10.1007/978-3-030-63322-6_56

[2] Farooq, U., Rahim, M.S.M., Sabir, N. et al. Advances in machine translation for sign language: approaches, limitations, and challenges. Neural Comput & Applic 33, 14357–14399 (2021). https://doi.org/10.1007/s00521-021-06079-3

[3] A. Ashrafi, V. S. Mokhnachev and A. E. Harlamenkov, "Improving Sign Language Recognition with Machine Learning and Artificial Intelligence," 2024 6th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE), Moscow, Russian Federation, 2024, pp. 1-6, doi: 10.1109/REEPE60449.2024.10479844.

[4] A. Ashrafi, V. S. Mokhnachev, Y. N. Philippovich, and L. P. Tsilenko, "Video classification using CNN-LSTM architecture for Bengali sign language" in Fundamental and applied sciences today XXVIII, p. 140-146. Pothi.com, Bengaluru, Karnataka, India (2022).

[5] A. Ashrafi, V. S. Mokhnachev, Y. N. Philippovich, A. Y. Harlamenkov, and L. P. Tsilenko, "Comparative Study Of Different Sign Language Recognition Methods", Natrazvitie, St. Petersburg, 2022, pp. 28-31.

[6] Ashrafi, A., Mokhnachev, V., Philippovich, Y., Harlamenkov, A., Chernenko, S. (2023). Russian Sign Language Recognition Using MediaPipe. In: Dolinina, O., et al. Artificial Intelligence in Models, Methods and Applications. AIES 2022. Studies in Systems, Decision and Control, vol 457. Springer, Cham. https://doi.org/10.1007/978-3-031-22938-1_21

[7] A. R. Várkonyi-Kóczy and B. Tusor, "Human–computer interaction for smart environment applications using fuzzy hand posture and gesture models", IEEE Trans. Instrum. Meas., 2011, 60, (5), pp. 1505– 1514.

[8] C. C. Hsieh and D. H. Liou, "Novel Haar features for real-time hand gesture recognition using SVM", Journal of Real-Time Image Process., 2015, 10, (2), pp. 357– 370.

[9] M. Tang, "Recognizing hand gestures with Microsoft's Kinect", Stanfordedu, 2011, 14, (4), pp. 303– 313

[10] Q. Chen, N. D. Georganas, and E. M. Petriu, "Hand gesture recognition using haarlike features and a stochastic context-free grammar", IEEE Trans. Instrum. Meas., 2008, 57, (8), pp. 1562– 1571.

[11] G. Plouffe and A. Cretu, "Static and dynamic hand gesture recognition in depth data using dynamic time warping", IEEE Trans. Instrum. Meas., 2016, 65, (2), pp. 305– 316.

[12] J. Han, L. Shao, and J. Shotton, "Enhanced computer vision with Microsoft Kinect sensor: a review", IEEE Trans. Cybern., 2013, 43, (5), pp. 1318– 1334.

[13] C. Wang, Z. Liu, and S. Chan, "Superpixel-based hand gesture recognition with Kinect depth camera", IEEE Trans. Multimed., 2015, 17, (1), pp. 29– 39.

[14] Y. Yao and Y. Fu, "Contour model-based hand-gesture recognition using the Kinect sensor", IEEE Trans. Circuits Syst. Video Technol., 2014, 24, (11), pp. 1935– 1944.

[15] F. Jiang, C. Wang, Y Gao, S Wu, and D Zhao, "Discriminating features learning in hand gesture classification", IET Comput. Vis., 2015, 9, (5), pp. 673– 680.

[16] Shuxian Liu, Wei Wang, Le Deng, Huan Xu, Cnn-trans model: A parallel dual-branch network for fundus image classification, Biomedical Signal Processing and Control, Volume 96, Part B, 2024, 106621, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2024.106621.

[17] Y. Farhan and A. Ait Madi, "Real-time Dynamic Sign Recognition using MediaPipe," 2022 IEEE 3rd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), Fez, Morocco, 2022, pp. 1-7, doi: 10.1109/ICECOCS55148.2022.9982822.

[18] Sign Language Detection and Recognition Using Media Pipe and Deep Learning Algorithm . (2024). International Journal of Scientific Research in Science and Technology, 11(2), 123-130. https://doi.org/10.32628/IJSRST52411223

[19] Laveen Chandwani, Jaydeep Khilari, Kunal Gurjar et al. Gesture based Sign Language Recognition system using Mediapipe, 27 June 2023, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-3106646/v1]

[20] J. Zamora-Mora and M. Chacón-Rivas, "Costa Rican Sign Language Recognition using MediaPipe," 2022 International Conference on Inclusive Technologies and Education (CONTIE), Cartago, Costa Rica, 2022, pp. 1-6, doi: 10.1109/CONTIE56301.2022.10004438.

[21] Zhang, Z., Cao, Y., Shi, X., & Liu, C. (2020). "Hand Gesture Recognition Using MediaPipe." International Journal of Advanced Computer Science and Applications, 11(6), 568-574. https://doi.org/10.14569/IJACSA.2020.0110672

[22] Alasmari A, Alhothali A, Allinjawi A. Hybrid machine learning approach for Arabic medical web page credibility assessment. Health Informatics Journal. 2022;28(1). doi:10.1177/14604582211070998

[23] Ashrafi, A., Mokhnachev, V.S., Philippovich, Y.N. Converting Video into Animation Using Video Tracking Technology for Developing a Multilingual Dictionary of Sign Language. IIASU'24 3rd All-Russian scientific conference "Artificial Intelligence in Automated Control Systems and Data Processing", within the framework of the congress "Russian Engineer" at Bauman Moscow State Technical University, 2024.

[24] Ashrafi A., Mokhnachev V.S. Conversion Of Indian Sign Language Into Text Using Gesture Recognition Method. Theory and Practice of Modern Science: the View of Youth. Материалы Всероссийской научно-практической конференции на английском языке. Научное издание. Сост. Е.Н. Лашина, М.С. Липатов. Под общей редакцией В.В. Кирилловой . Санкт-Петербург, 2022. С. 204-211.

[25] Ashrafi, A., Mokhnachev, V.S. (2023) Designing User-Friendly Interfaces For A Multilingual Sign Language Dictionary, Artificial Intelligence in Automated Control and Data Processing Systems, Collection of Articles of the 2nd All-Russian Scientific Conference: In 5 Volumes, April 27–28, Moscow.