



Data Article

NSL23 dataset for alphabets of Nepali sign language

Jhuma Sunuwar, Samarjeet Borah*, Aditi Kharga

Sikkim Manipal Institute of Technology, Sikkim Manipal University, Sikkim, 737136, India

ARTICLE INFO

Article history:

Received 16 April 2023

Revised 16 January 2024

Accepted 16 January 2024

Available online 23 January 2024

Dataset link: [Nepali Sign Language -Consonant and Vowel \(Original data\)](#)

Keywords:

Nepali sign language

NSL

NSL23

Consonant

Vowel

Sign language translation

Computer vision

ABSTRACT

Nepali Sign Language (NSL) is used by the Nepali-speaking community in Nepal and in Indian states such as Sikkim, the hilly region of North Bengal, some parts of Uttarakhand, Meghalaya, and Assam. It consists of the International Manual Alphabet (A-Z), Nepali consonants, vowels, conjunct letters, and numbers represented in the form of one-handed fingerspelling or Nepali manual alphabet. The standard gestures for NSL have been published by the Nepal National Federation of the Deaf & Hard of Hearing (NFDH).

To learn Nepali Sign Language, the first step is to understand its alphabet set. The use of technology can help ease the learning process. One of the application areas of computer vision is translating sign language gestures to either text or audio to facilitate communication. This is an open research area. However, NSL translation is one of the less explored research areas because there is no dataset available to work on for NSL. This paper introduces the Nepali Sign Language Dataset (NSL23), which is the first of its kind and includes vowels and consonants of the Nepali Sign Language alphabet. The dataset consists of .mov videos performed by 14 volunteers who have demonstrated 36 consonant signs and 13 vowel signs either in one full video or character by character. The dataset has been prepared under various conditions, including normal lighting, dark lighting conditions, prepared environments, unprepared environments, and real-world environments. The volunteers who performed the NSL gesture have been classified as 9 beginners who are using NSL for the first time and 5 experts who have been using NSL for 5 to

* Corresponding author.

E-mail address: samarjeet.b@smit.smu.edu.in (S. Borah).

25 years. NSL23 contains 630 total videos representing 1205 gestures. The dataset can be used to train machine learning models to classify the alphabet set of NSL and further develop a sign language translator.

© 2024 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

Specifications Table

Subject	Computer Science
Specific subject area	Hand Gesture Detection, Gesture Categorization, Sign Language Detection
Type of data	Video (.mov)
How the data were acquired	<p>Data was acquired using Nikon D3200 and Iphone7 Plus. Configuration for both is given below:</p> <ul style="list-style-type: none">• Nikon D3200 24.2MP Digital SLR Camera, CMOS Sensor, Video Resolution-1280×720,• Mobile phone camera Iphone7 Plus: 12 MP, f/1.8, 28 mm (wide), 1/3", PDAF, OIS, 12 MP, f/2.8, 56 mm (telephoto), 1/3.6", AF, 2x optical zoom. Video: 4K@30fps, 1080p@30/60/120fps, 720p@240fps.
	<p>Data acquisition Process</p> <ul style="list-style-type: none">• Data were acquired in the form of video(.mov)• Each volunteer was asked to perform hand gestures facing the camera. Gesture was performed using only the right hand.• To ease cropping and extraction operation during pre-processing, the hand performing the gesture was placed beside the face.• Duration of video captured character by character varied from 1 to 5 s and video comprising all character comprises of 14.45 s to 4.5 min.
	<p>Environment- Prepared, unprepared, Lighting- Bright, Dark, NormalLocation-Indoor, Outdoor has been considered while capture the data.</p> <ul style="list-style-type: none">• Each volunteer has performed 36 Consonant and 13 vowels either as a full video or one character per video.• Repetition of data is done to generate cropped video in the case of volunteer number S1, S2 and S3.• Volunteer number S3, S4, S5, S6 and S14 have performed all characters in single shot video as they are experienced users of NSL.
Data format	Raw
Description of data collection	<p>Data was collected from 14 volunteers, categorized as 5 Native (>25 years and >5 years of using NSL) and 9 beginners learning NSL.Dataset has 630 total videos containing 1205 gestures for 36 consonants and 13 vowels of NSL. 516 Videos for consonants that have 971 gestures, 114 videos for vowel that has 234 gestures.Data was collected in two types of environments.</p> <p>(1) Prepared environment - Black background by putting a black chart paper at the background of the wall where the gesture is performed.</p> <p>(2) Unprepared environment- No black chart paper was used to change the background while performing the gesture.</p> <p>Real world environment indoor and outdoor has been considered.Lighting condition – Natural lighting condition has been considered while capturing the data at indoor and outdoor i.e., no extra light was focused on the volunteer performing the gesture. Bright and Dark condition for indoor videos is referring the videos captured by turning the room lights and off respectively.</p> <p>(continued on next page)</p>

Data source location	<ul style="list-style-type: none">• Institution: Sikkim Manipal Institute of Technology, Sikkim Manipal University• City/Town/Region: Rangpo/Majitar• Country: India• Latitude and longitude: 27°10'57.5"N 88°30'07.8"E
Data accessibility	<p>Repository name: Nepali Sign Language - Consonant and Vowel</p> <p>Data identification number: 10.5281/zenodo.10215687 (version1)</p> <p>10.5281/zenodo.10478554 (version2)</p> <p>Direct URL to data: https://shorturl.at/lsSUW</p>

1. Value of the data

- The users of Nepali Sign Language will greatly benefit out of this dataset as it is first of its kind and will be used by researchers to develop automatic Nepali sign language recognition application. There are many existing datasets for other languages [2–5,7,8,10–13,15,17,19–23]
- This dataset can be used to develop Nepali sign language translator a mean of communication between the deaf and non-deaf communities.
- This dataset will help researchers and developers to train and test machine learning, deep learning models [3,4,18] created specifically for automated Nepali Sign Language hand signs detection and recognition.
- The Nepali sign language dataset opens up a new avenue for future study and development of real-world gesture recognition for researchers as it comprises of real-world condition [3,6,8] where varying lighting conditions, backgrounds, and hand positioning has been considered.
- This will benefit Nepali speakers who want to learn Nepali Sign Language as well as those developing translation applications to serve as a communication tool between NSL user and non-user [1,14,19].
- It can be used to create new similar dataset or to expand the dataset by replicating the samples or adding new samples of Nepali sing language in different background conditions, lighting conditions, and orientation to further improve NSL23 dataset.

2. Background

Several hand gesture datasets are available for public use [2–5,7,8,10–13,15,17,19–23]. However, most of them lack variations in the gestures they provide. While some sign language datasets that are publicly available offer a collection of frames to represent a gesture, others provide videos that were collected under the same environmental and lighting conditions. The paper aims to create a comprehensive dataset that includes various lighting conditions such as indoor, outdoor, bright, dark, and natural lighting. The dataset will also consider prepared and unprepared backgrounds. The scenario for the dataset involves volunteers performing gestures at different positions and heights in a real-world setting. The objectives of the dataset are to:

- To create a dataset that provides a Nepali Sign Language alphabet set (Consonant and Vowel) and make it available publicly.
- To provide a dataset having variations in environments and lighting conditions.
- To provide the dataset in raw form so that users can preprocess the data as per their requirement and use it to train supervised or semi-supervised or test the supervised and unsupervised machine learning model and deep learning models [6,9,10,16,18].
- To encourage researchers to start working in Nepali Sign Language translation.

3. Data Description

The NSL23 dataset contains two folders: one for consonants and one for vowels. Each folder contains subfolder which further contains videos. The details of both folders are provided in Tables 1 and 2 respectively.

Table 1
Description of NSL23 Consonant Dataset.

Sl No	Name of the folder	No. of videos	Description
1	S1_NSL_Consonant_Dark	36	Indoor unprepared in dark condition captured using phone camera
2	S1_NSL_Consonant_Bright	36	Indoor unprepared in normal condition captured using phone camera
3	S1_NSL_Consonant_Dark_Cropped	36	Indoor unprepared in dark condition captured using phone camera
4	S1_NSL_Consonant_Bright_Cropped	36	Indoor unprepared in normal condition captured using phone camera
5	S2_NSL_Consonant_Dark	36	Indoor unprepared in dark condition captured using phone camera
6	S2_NSL_Consonant_Bright	36	Indoor unprepared in normal condition captured using phone camera
7	S2_NSL_Consonant_Dark_Cropped	36	Indoor unprepared in dark condition, preprocessed video captured using phone camera
8	S2_NSL_Consonant_Bright_Cropped	36	Indoor unprepared in normal condition, preprocessed video captured using phone camera
9	S3_NSL_Consonant_Prepared	2 [2×36=72 gestures]	2 indoor prepared black background captured using camera and phone camera
10	S3_NSL_Consonant_Prepared_Cropped	36	Indoor prepared black background captured using phone camera
11	S3_NSL_Consonant_Real_World	2 [2×36=72gestures]	2 outdoor unprepared environment-real worlds captured using phone camera
12	S4_NSL_Consonant_Prepared	2 [2×36=72 gestures]	Indoor prepared black background taken from mobile camera and phone camera in slightly different angle.
13	S5_NSL_Consonant_Prepared	1 [36 gesture]	Indoor prepared black background captured using camera
14	S6_NSL_Consonant_Prepared	2 [2×36=72 gestures]	Indoor prepared black background taken from mobile camera and phone camera in slightly different angle
15	S6_NSL_Consonant_RealWorld	1 [36 gestures]	Indoor unprepared environment captured using phone camera
16	S7_NSL_Consonant_Prepared	36	Indoor prepared black background captured using phone camera
17	S8_NSL_Consonant_Prepared	36	Indoor prepared black background captured using phone camera
16	S9_NSL_Consonant_Prepared	1 [36 gestures]	Indoor prepared black background captured using phone camera
17	S10_NSL_Consonant_Prepared	36	Indoor prepared black background captured using phone camera
18	S11_NSL_Consonant_Prepared	36	Indoor prepared black background captured using phone camera
19	S12_NSL_Consonant_Prepared	35	Indoor prepared black background captured using phone camera
20	S13_NSL_Consonant_Prepared	1 [36 gestures]	Indoor prepared black background captured using phone camera
21	S14_NSL_Consonant_RealWorld	1 [36 gestures]	Indoor unprepared background captured using phone camera

Table 2

Description of NSL23 Vowel Dataset.

Sl No	Name of the folder	No. of videos	Description
1	S1_NSL_Vowel_Unprepared_Bright	13	Indoor Unprepared Background captured using phone camera
2	S1_NSL_Vowel_Unprepared_Bright_Cropped	13	Indoor Unprepared Background captured using phone camera
3	S1_NSL_Vowel_Unprepared_Dark	13	Indoor Unprepared Background captured using phone camera
4	S1_NSL_Vowel_Unprepared_Dark_Cropped	13	Indoor Unprepared Background captured using phone camera
5	S2_NSL_Vowel_Unprepared_Bright	13	Indoor Unprepared Background captured using phone camera
6	S2_NSL_Vowel_Unprepared_Bright_Cropped	13	Indoor Unprepared Background captured using phone camera
7	S2_NSL_Vowel_Unprepared_Dark	13	Indoor Unprepared Background captured using phone camera
8	S2_NSL_Vowel_Unprepared_Dark_Cropped	13	Indoor Unprepared Background captured using phone camera
9	S3_NSL_Vowel_Prepared	2 [2×13=26 gestures]	Indoor Prepared Background captured using camera and phone camera
10	S3_NSL_Vowel_Unprepared_RealWorld	1 [13 gesture]	Outdoor Unprepared Background captured using phone camera
11	S4_NSL_Vowel_Prepared	2 [2×13=26 gestures]	Indoor Prepared Background captured using camera and phone camera
12	S5_NSL_Vowel_Prepared	1 [13 gestures]	Indoor Prepared Background captured using camera
13	S6_NSL_Vowel_Prepared	2 [2×13 =26 gestures]	Indoor Prepared Background captured using camera and phone camera
14	S6_NSL_Vowel_Unprepared	1 [13 gestures]	Indoor Unprepared Background captured using phone camera
15	S14_NSL_Vowel_Unprepared	1 [13 gestures]	Indoor Unprepared Background captured using phone camera

Table 1 shows the organization of subfolders which contains 36 consonant of Nepali sign language.

Table 2 shows the organization of subfolders which contains 13 vowels of Nepali sign language.

Tables 3 and 4 to identify each video alphabetically. The direction of the arrow in the hand gesture represents the direction of movement of the hand while performing the dynamic gesture. Volunteers were given the table to understand and learn the gesture before capturing their video as shown in Fig. 3.







The dataset has been made publically available for use in Zenodo platform.Using the link : <https://shorturl.at/lsSUW>, user will be able to download the following files:

- i. Details Dataset.docx – Gives description table for each video
- ii. NSL_Consonant_Part_1.zip
- iii. NSL_Consonant_Part_2.zip
- iv. NSL_Consonant_Part_3.zip
- v. NSL_Vowel.zip

NSL_Consonant

Each folder inside NSL_Consonant and NSL_vowel has been labeled using the naming convention as follows: volunteer number (S1 to S14) _ NSL _ Consonant (or Vowel) _Bright (or Dark, Prepared, Unprepared, RealWorld) _Cropped (if cropped video). Each folder consists of videos in which volunteers have demonstrate the hand gestures for consonant and vowel alphabets of NSL respectively. These videos are named using the naming convention specified in Table 3 and Table 4. If a video displays one character of NSL consonant/Vowel at a time, otherwise the word "all" is used to signify that the video contains all the alphabet of NSL consonant/vowel in a single video. Table 1 and Table 2 provide the detail content of NSL_Consonant and NSL_Vowel

Table 3
Nomenclature used for consonant videos (All Static Gesture).














Srl. No	Nomenclature used	Nepali Alphabet	Nepali Sign Language	Srl. No	Nomenclature used	Nepali Alphabet	Nepali Sign Language
1	KA	क		13	DA	ड	
2	KHA	ख		14	DHA	ढ	
3	GA	ग		15	NA	ण	
4	GHA	घ		16	TAA	त	
5	NGA	ङ		17	THAA	थ	
6	CHA	च		18	DAA	ढ	
7	CHHA	छ		19	DHAA	ध	
8	JA	ज		20	NAA	ण	
9	JHA	झ		21	PA	प	
10	YAN	य		22	PHA	फ	
11	TA	ट		23	BA	ब	
12	THA	ठ		24	BHA	भ	

(continued on next page)

Table 3 (continued)

25	MA	म		31	M_SHA	ष	
26	YA	य		32	D_SHA	स	
27	RA	र		33	HA	ह	
28	LA	ल		34	KSHA	क्ष	
29	WA	व		35	TRA	त्र	
30	T_SHA	श		36	GYA	ज्ञ	

Table 4
Nomenclature used for vowel videos (Static and Dynamic gesture).

Srl. No	Nomenclature used	Nepali Alphabet	Nepali Sign Language	Srl. No	Nomenclature used	Nepali Alphabet	Nepali Sign Language
1	A	अ		8	E	ए	
2	AA	आ		9	AI	ऐ	
3	I	इ		10	O	ओ	
4	II	ई		11	AU	औ	
5	U	उ		12	AM	अं	
6	UU	ऊ		13	AH	अः	
7	RI	ऋ					

respectively. Users can follow the above given naming convention to access the corresponding data as per requirement.

4. Experimental Design, Materials, and Methods

To learn any sign language [2–4,7,20,22], one must start with its alphabets. This paper focuses on the construction of Nepali Sign Language alphabets (consonant and vowel) and has been named NSL23. To start with the data set preparation that can be used by Machine learning model to recognize real world NSL gesture, training data needs to be prepared. A faculty of a special school was approached to assist with the data collection process. He teaches Nepali sign language to differently able students of class 1 to class 5. With his support it was possible to collect raw data in the form of video in a prepared and un-prepared environment. The prepared environment was created by pasting black chart paper at the background while the volunteer performing the gesture pose facing the camera kept in front of them. Fig. 1 shows some volunteers and the environment used during data collection. Total there are 14 volunteers out of whom 5 volunteers are native users of Nepali sign language and are either the faculty using NSL or students who have been using it for more than 5 years.

Data has been captured in first persons’ view [8] which means the camera is placed in front of the volunteer performing the gesture. To increase the scope of the dataset various environmental and illumination conditions have been considered. The videos are taken indoor- prepared as shown in Fig. 1, unprepared, and outdoor to make it more dynamic. Fig. 2 show the different real world environment considered while acquiring the data.



Fig. 1. Volunteers posing for NSL23 dataset in prepared environment.

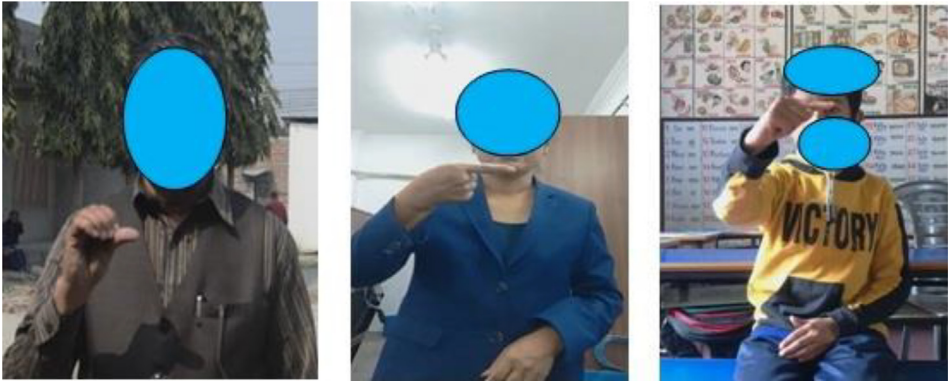


Fig. 2. Outdoor and Indoor Environment, Unprepared Environment to Showcase Real World Scenario.

Initial phase of collecting NSL23 dataset comprises of 630 total videos, 1205 gestures performed by 14 volunteers labeled as S1 to S14 to identify each volunteer. Volunteers S3, S4, S5, S6, and S14 are experts in NSL, so they provided consonant and vowel gestures in one take. However, the remaining volunteers were new to NSL, so their signs had to be captured character by character as shown in Fig. 3. A prepared environment was created by putting black chart paper in the background to ease the segmentation process of separating foreground and background. For the dataset, outdoor videos were shot in a natural light environment, while indoor videos were shot with natural light as well as by turning on room light. To illustrate the illumination condition, all the videos have been labeled as dark or bright depending on the lighting environment used during the video shoot. To increase variation in the dataset, some videos were preprocessed using Adobe Express tools to crop and keep only the hand portion. These videos are labeled as “cropped”.

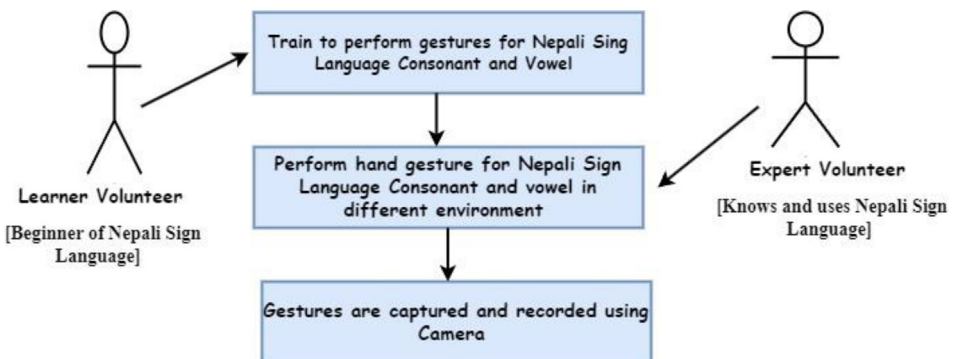


Fig. 3. Process prior to data generation.



Fig. 6. Indoor Unprepared Background (Bright lighted condition).

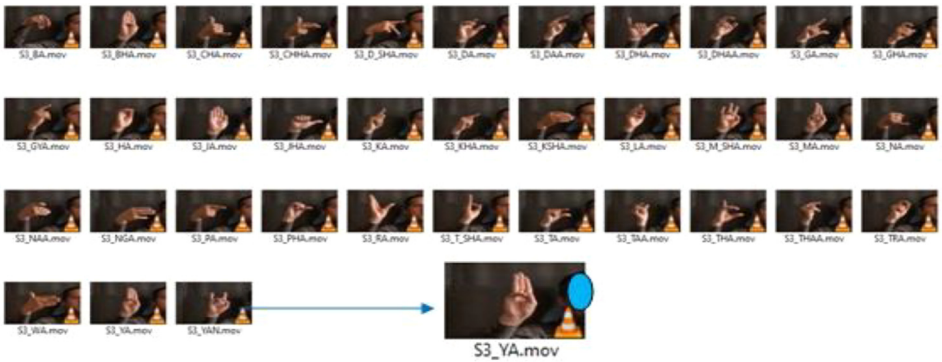


Fig. 7. Indoor Prepared Black Background and cropped.

Video containing incomplete or wrong gesture was manually discarded. The videos were trimmed and cropped using the Adobe Express (online video editing website) [source: <https://www.adobe.com/express>] as shown in Fig. 7 for videos labeled as cropped. Other preprocessing approach has not been performed in NSL23 to encourage users apply preprocessing techniques as per their need to make the dataset suitable for their purpose.

Ethics Statements

This data collection work has involvement of human subjects. Consents from all the subjects/participants are taken in writing form and scanned copy of the same is attached as supplementary material with this submission.

Data Availability

Nepali Sign Language -Consonant and Vowel (Original data) (zenodo.org)

CRediT Author Statement

Jhuma Sunuwar: Data curation, Formal analysis, Writing – original draft; **Samarjeet Borah:** Conceptualization, Project administration, Supervision, Writing – review & editing; **Aditi Kharga:** Data curation.

Acknowledgments

The work of data collection would not be possible without the help of Mr. Hari Prasad Adhikari teacher at Himali Secondary School, Laxmipur, Jhapa, Nepal for granting permission to visit his school, provide help in understanding Nepali Sign Language and give consent to collect data. In addition, Mr. Adhikari introduced Ms. Gita Shrestha, Sign Language teachers of Himali Secondary School and two students Marmik Bhattarai, and Mira Rajbanshi who have also provided the consent to collect data. Mrs. Shilu Sharma, Nepali Sign Language News Presenter at Nepal Television and who has also worked at Nepal Television and National Deaf Federation-Nepal has provided the standard published NSL alphabet set template referred while collecting data. Finally, we would like to acknowledge volunteers of Sikkim Mr. Kamal Subba, Mr. Pravesh Adhikari, Miss. Samten Lhamu Bhutia, Miss. Sitoshna Chhetri, Miss. Mahima Rai, Mr. Manhang and Mrs. Kapila Sharma from giving their valuable time to learn and pose NSL alphabet set. Authors are thankful to Sikkim Manipal Institute of Technology for providing the platform to do research in the area of computer vision, helping to develop NSL23 dataset.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] V. Adithya, R. Rajesh, Hand gestures for emergency situations: a video dataset based on words from Indian sign language, *Data Brief* 31 (2020) 106016.
- [2] A.B.H. Amor, O. El Ghoul, M. Jemni, An EMG dataset for Arabic sign language alphabet letters and numbers, *Data Brief* (2023) 109770.
- [3] J. Bora, S. Dehingia, A. Boruah, A.A. Chetia, D. Gogoi, Real-time assamese sign language recognition using MediaPipe and deep learning, *Procedia Comput. Sci.* 218 (2023) 1384–1393.
- [4] S. Das, M.S. Imtiaz, N.H. Neom, N. Siddique, H. Wang, A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier, *Expert Syst. Appl.* 213 (2023) 118914.
- [5] T.F. Dima, M.E. Ahmed, Using YOLOv5 algorithm to detect and recognize American sign language, in: 2021 International Conference on Information Technology (ICIT), IEEE, 2021, pp. 603–607.
- [6] B. Garcia, S.A. Viesca, Real-time American sign language recognition with convolutional neural networks, *Convolutional Neural Netw. Vis. Recognit.* 2 (2016) 225–232.
- [7] A. Hasib, J.F. Eva, S.S. Khan, M.N. Khatun, A. Haque, N. Shahrin, ... M.R. Hussein, BDSL 49: a comprehensive dataset of Bangla sign language, *Data Brief* (2023) 109329.
- [8] V.T. Hoang, HGM-4: a new multi-cameras dataset for hand gesture recognition, *Data Brief* 30 (2020) 105676.
- [9] J. Huang, W. Zhou, H. Li, W. Li, Sign language recognition using 3D convolutional neural networks, in: 2015 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2015, pp. 1–6.
- [10] A. Imran, A. Razaq, I.A. Baig, A. Hussain, S. Shahid, T.U. Rehman, Dataset of Pakistan sign language and automatic recognition of hand configuration of Urdu alphabet through machine learning, *Data Brief* 36 (2021) 107021.
- [11] M.M. Islam, M.R. Uddin, M.J. Ferdous, S. Akter, M.N. Akhtar, BdSLW-11: dataset of Bangladeshi sign language words for recognizing 11 daily useful BdSL words, *Data Brief* 45 (2022) 108747.
- [12] A.A.J. Jim, I. Rafi, M.Z. Akon, U. Biswas, A.A. Nahid, KU-BdSL: an open dataset for Bengali sign language recognition, *Data Brief* (2023) 109797.
- [13] R.T. Johari, R. Ramli, Z. Zulkoffli, N. Saibani, MyWSL: Malaysian words sign language dataset, *Data Brief* (2023) 109338.
- [14] S. Kim, Y. Ji, K.B. Lee, An effective sign language learning with object detection-based ROI segmentation, in: 2018 Second IEEE International Conference on Robotic Computing (IRC), IEEE, 2018, pp. 330–333.
- [15] G. Latif, N. Mohammad, J. Alghazo, R. AlKhalaf, R. AlKhalaf, ArASL: Arabic alphabets sign language dataset, *Data Brief* 23 (2019) 103777.
- [16] T.H. Le, D.W. Jaw, I.C. Lin, H.B. Liu, S.C. Huang, An efficient hand detection method based on convolutional neural network, in: 2018 7th International Symposium on Next Generation Electronics (ISNE), IEEE, 2018, pp. 1–2.
- [17] Mavi, A., & Dikle, Z. (2022). A new 27 class sign language dataset collected from 173 individuals. *arXiv preprint arXiv:2203.03859*.
- [18] A. Mujahid, M.J. Awan, A. Yasin, M.A. Mohammed, R. Damaševičius, R. Maskeliūnas, K.H. Abdulkareem, Real-time hand gesture recognition based on deep learning YOLOv3 model, *Appl. Sci.* 11 (9) (2021) 4164.

- [19] C. Nuzzi, S. Pasinetti, R. Pagani, G. Coffetti, G. Sansoni, HANDS: an RGB-D dataset of static hand-gestures for human-robot interaction, *Data Brief* 35 (2021) 106791.
- [20] Özdemir, O., Kindiroğlu, A. A., Camgöz, N. C., & Akarun, L. (2020). Bosporussign22k sign language recognition dataset. arXiv preprint arXiv:2004.01283.
- [21] A.A.I. Sidig, H. Luqman, S. Mahmoud, M. Mohandes, KArSL: Arabic sign language database, *ACM Trans. Asian Low-Resource Lang. Inf. Process. (TALLIP)* 20 (1) (2021) 1–19.
- [22] S.R. Yeduri, D.S. Breland, S.B. Skriubakken, O.J. Pandey, L.R. Cenkeramaddi, Low resolution thermal imaging dataset of sign language digits, *Data Brief* 41 (2022) 107977.
- [23] Y. Zhang, C. Cao, J. Cheng, H. Lu, Egogesture: a new dataset and benchmark for egocentric hand gesture recognition, *IEEE Trans. Multimed.* 20 (5) (2018) 1038–1050.