

# Sign-Language Detection

Subjects: **Computer Science**, **Artificial Intelligence**

Contributor: Nasima Begum , Rashik Rahman , Nusrat Jahan , Saqib Sizan Khan , Tanjina Helaly , Ashraful Haque , Nipa Khatun

Sign language is the most commonly used form of communication for persons with disabilities who have hearing or speech difficulties. However, persons without hearing impairment cannot understand these signs in many cases. As a consequence, persons with disabilities experience difficulties while expressing their emotions or needs. Thus, a sign character detection and text generation system is necessary to mitigate this issue.

computer vision

Sign Language Recognition

sign-language detection

CNN

dataset

## 1. Introduction

The most-widespread communication means for persons with hearing or speech disabilities is sign language. Bengali Sign Language (BSL) is composed of precise hand motions produced with one or two hands, where the arrangement of the fingers and the position of the palm together form the characters of the language <sup>[1][2]</sup>. The signs that are used to represent Bengali characters and numerals are called BSL. In Bangladesh, there are over 1.7 million people who have hearing impairment or speech impairment. These hearing- or speech-impaired individuals use BSL for communication. However, due to the lack of knowledge of sign language, persons with such disabilities cannot convey their opinions and emotions properly. To address this issue, an interpreter system is needed that can detect Bengali signs characters and translate them into Bengali text. It can reduce the communication gap between person with and without disabilities.

Currently, most research emphasizes vision-based <sup>[3]</sup> or sensor-based <sup>[4]</sup> systems. However, further research needs to be done in this domain to introduce an efficient and convenient system for sign-language interpretation. At present, various benchmarking datasets for BdSL detection can be found, and there has been much research on the subject. However, these datasets are insufficient for developing and testing deep learning models, and most of them are closed-source. As a result, it becomes difficult to train Convolutional Neural Network (CNN) models to detect BSL.

Sign-language detection is one of the challenging tasks in computer vision. Numerous related works <sup>[5][6][7][8][9]</sup> on detecting the individual characters of Bangla Sign Language can be found. However, hardly any work has been done with generating text from sign images. Many research studies only included around 25–35 sign characters for their detection models. The dataset we have used, has 49 different characters. Additionally, researchers use various techniques, i.e., K-Nearest Neighbors (KNN), Support Vector Machine (SVM), shallow neural networks, and image processing, for detection purposes. However, these models are trained on a small dataset consisting of

only 2000–3000 pictures in total. Furthermore, previous works not only used a smaller amount of data but also a smaller number of classes. The maximum number of classes was 36, which was proposed by Hoque et al. [10]. The YoloV4 [11] detection model is quite popular among researchers for detecting sign language [12][13]. Researchers [13] developed a BSL detection model using YOLOv4. In addition to detecting or localizing the region of the waved sign in an image or video frame, the YoloV4 model also predicts the class or label of which sign is being waved. In a nutshell, this detection model also recognizes the localized area of the hand sign.

## 2. Sign Language Recognition

Sign Language Recognition (SLR) is a significant computer vision task that has been extensively studied. The prevalent SLR approaches can be categorized into two groups: recognizing a single sign at a time [14] and recognizing multiple signs from video content input [15]. However, there is lack of sufficient study for Bangla SLR and/or detection. This section discusses previous work related to Bangla SLR and detection.

Venugopalan et al. [16] proposed a dataset of commonly used Indian Sign Language (ISL) words for hearing-impaired COVID patients in emergency cases. They also presented a deep convolutional LSTM hybrid model for recognizing the proposed ISL words and acquired an accuracy of 83.36%. The authors in [17] developed a system that can recognize Arabic hand signs of 31 classes and translates them into Arabic speech. The system provides 90% for recognizing Arabic hand signs. The authors propose a deep multi-layered CNN model in [8] to detect 36 static English signs, including digits and alphabet characters and 26 dynamic English signs such as emotional hand gestures. Their proposed model achieved 99.89% accuracy in the blind testing phase. One drawback of this model is that it cannot detect signs in real time. Khan et al. [18] proposed a grammar-based translation model to translate English sentences into equivalent Pakistani Sign Language and got a Bilingual Evaluation Understudy (BLEU) score of 0.78. However, the model is unable to translate compound and complex sentences correctly.

For Bangla, OkkhorNama is an image dataset is proposed in [19] for real-time object detection. It contains over 12,000 images of 46 classes. The dataset is used to train different versions of YOLOv5 for performance analysis. However, the number of classes is not significant. Afterwards, a dataset named BdSL36, which includes 26,713 images of 36 different classes, was proposed by [10]. This dataset is further divided into two sections. The first section contains images for developing a detection model, and the second section contains images for developing a recognition model. To evaluate their dataset, the authors trained ResNet-50 on their dataset and achieved an accuracy of 98.6%. Although the number of images in the dataset is significantly large due to data augmentation, background removal, affine transformation, and many other methods applied, it only has 36 classes. The authors of [20] developed a dataset named Shongket, which consists of 36 classes of distinct Bengali signs. They also trained a few CNN models using their dataset and achieved about 95% accuracy. The dataset BDSLInfinite, consisting of 37 different signs, was developed by Urme et al. [6] and contains 2000 images, and a trained recognition model using Xception architecture achieved 98.93% accuracy as well as 48.53ms response time.

The authors of [21] proposed a custom CNN architecture and trained it on a dataset that consists of 100 different static sign classes. On colored and grayscale images, the proposed method obtained training accuracy rates of

99.72% and 99.90%, respectively. Basnin et al. [22] proposed an integrated CNN-LSTM model for Bangla Lexical Sign Language Recognition using a dataset that contains 13,400 images of 36 distinct classes of Bangla lexical signs. Their proposed model achieved 90% training accuracy and 88.5% testing accuracy. Moreover, they used high computational techniques for data pre-processing, which made their model computationally expensive.

An application that speaks the results in Bangla after automatically detecting hand-sign-based digits was developed by the authors of [23]. They utilized deep neural networks to complete the task, and their proposed model achieved an accuracy of 92%. Islam et al. [24] used four different finely tuned transfer learning methods to recognize 11 different Bengali sign words and got an accuracy of 98%. Shurid et al. [25] proposed a Bangla sign language recognition and sentence generation model and achieved 90% accuracy using their augmented dataset. However, their proposed model could not work correctly to recognize critical sign gestures. Angona et al. [26] developed a computer system to recognize 36 different classes of BSL and translated them into text format. For sign recognition, they used the MobileNet version 1 model and got an accuracy of 95.71%. Podder et al. [27] designed a sign language transformer that can aid in establishing communication with doctors remotely by translating sign language into text and speech. It also translates the speeches of doctors into sign language. However, the system is not yet tested for people with disabilities in Bangladesh. A Scale-Invariant Feature Transform (SIFT) technique and Convolutional Neural Network (CNN) are used in the proposed system of [7] to recognize one-handed gestures of 38 Bangla Signs. However, the used dataset is relatively low. Rafiq et al. [9] proposed a translation system that can translate different Bengali word signs into Bengali speech using a seven-layered custom sequential CNN model. The system trained with 1500 images of 10 different word signs and achieved 97% test accuracy as well as an average response time of 120.6ms.

The authors in [5] proposed a method to detect and recognize American Sign Language using the YOLOv5 algorithm and achieved an overall F1-score of 0.98. However, their dataset contains only 36 different sign classes. The authors of [13] proposed a real-time Bangla sign-language detection model and generated textual and audio speech. They created a dataset of 49 different classes containing 12,500 images of BDSL. They used Yolov4 for detecting hand gestures and got an overall accuracy of 97.95%. However, while generating some common words, the system faces some difficulties. A rule-based system is proposed by the authors of [28] to interpret Bengali text and voice speech into BSL. The system achieved an accuracy of 96.03% for voice interpretation and 100% for text translation. However, the system was trained and tested with only Bangla numerals. Khan et al. [29] proposed a CNN and customized Region-of-Interest (ROI) segmentation-based BDSL translator device that can translate only five sign gestures. It demonstrates about 94% accuracy in detecting signs in real-time. Das et al. [30] proposed a hybrid model for automatic recognition of Bangla Sign Language (BSL) numerals and alphabet characters using a deep transfer learning-based convolutional neural network with a random forest classifier. The proposed system is verified on 'IsharaBochon' and 'Ishara-Lipi' datasets and achieves 91.67% and 97.33% accuracy, respectively, for both character and digit recognition. The authors of [31] proposed a method that includes three steps: segmentation, augmentation, and CNN-based classification, and is evaluated on three benchmark datasets. The segmentation approach accurately identifies gesture signs using a concatenated approach with YCbCr, HSV, and a watershed algorithm. A CNN-based model called BenSignNet is applied to extract features and for classification. Hassan et al. [32] present a low-cost Bangla sign language recognition model that uses neural networks to convert

signs into Bangla text. The dataset was manually captured, and image processing techniques were used to map actions to relevant text. The system achieves an accuracy of approximately 92% and can be used for applications such as sign language tutorials or dictionaries. Akash et al. [33] proposed a real-time Bangla Sign Language (BdSL) detection system that aims to generate Bangla sentences from a sequence of images or a video feed. The system uses the Blazepose algorithm to detect sign-language body postures and a Long Short-Term Memory (LSTM) network to train the data. After training for 85 epochs, the model achieved a training accuracy of 93.85% and a validation accuracy of 87.14%. The authors of [34] are working on isolated sign language recognition using PyTorch and YOLOv5 for video classification. The system aims to help people with hearing and speech disabilities to interact with society. The authors achieved an accuracy rate of 76.29% on the training dataset and 51.44% on the testing dataset.

## References

1. Sanzidul Islam, M.; Sultana Sharmin Mousumi, S.; Jessan, N.A.; Shahariar Azad Rabby, A.; Akhter Hossain, S. Ishara-Lipi: The First Complete Multipurpose Open Access Dataset of Isolated Characters for Bangla Sign Language. In Proceedings of the 2018 International Conference on Bangla Speech and Language Processing (ICBSLP), Sylhet, Bangladesh, 21–22 September 2018; pp. 1–4.
2. Rahaman, M.A.; Jasim, M.; Ali, M.H.; Hasanuzzaman, M. Bangla language modeling algorithm for automatic recognition of hand-sign-spelled Bangla sign language. *Front. Comput. Sci.* 2020, 14, 143302.
3. Kudrinko, K.; Flavin, E.; Zhu, X.; Li, Q. Wearable sensor-based sign language recognition: A comprehensive review. *IEEE Rev. Biomed. Eng.* 2020, 14, 82–97.
4. Sharma, S.; Singh, S. Vision-based sign language recognition system: A Comprehensive Review. In Proceedings of the 2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 26–28 February 2020; IEEE: New York, NY, USA, 2020; pp. 140–144.
5. Dima, T.F.; Ahmed, M.E. Using YOLOv5 Algorithm to Detect and Recognize American Sign Language. In Proceedings of the 2021 International Conference on Information Technology (ICIT), Amman, Jordan, 14–15 July 2021; IEEE: New York, NY, USA, 2021; pp. 603–607.
6. Urmee, P.P.; Al Mashud, M.A.; Akter, J.; Jameel, A.S.M.M.; Islam, S. Real-time bangla sign language detection using xception model with augmented dataset. In Proceedings of the 2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), Bangalore, India, 15–16 November 2019; IEEE: New York, NY, USA, 2019; pp. 1–5.
7. Shanta, S.S.; Anwar, S.T.; Kabir, M.R. Bangla sign language detection using sift and cnn. In Proceedings of the 2018 9th International Conference on Computing, Communication and

- Networking Technologies (ICCCNT), Bengaluru, India, 10–12 July 2018; IEEE: New York, NY, USA, 2018; pp. 1–6.
8. Bhadra, R.; Kar, S. Sign Language Detection from Hand Gesture Images using Deep Multi-layered Convolution Neural Network. In Proceedings of the 2021 IEEE Second International Conference on Control, Measurement and Instrumentation (CMI), Kolkata, India, 8–10 January 2021; IEEE: New York, NY, USA, 2021; pp. 196–200.
  9. Rafiq, R.B.; Hakim, S.A.; Tabashum, T. Real-time Vision-based Bangla Sign Language Detection using Convolutional Neural Network. In Proceedings of the 2021 International Conference on Advances in Computing and Communications (ICACC), Kochi, India, 21–23 October 2021; IEEE: New York, NY, USA, 2021; pp. 1–5.
  10. Hoque, O.B.; Jubair, M.I.; Akash, A.F.; Islam, S. Bdsl36: A dataset for bangladeshi sign letters recognition. In Proceedings of the 15th Asian Conference on Computer Vision, Kyoto, Japan, 30 November–4 December 2020.
  11. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. arXiv 2020, arXiv:2004.10934.
  12. Ma, D.; Hirota, K.; Dai, Y.; Jia, Z. Dynamic Sign Language Recognition Based on Improved Residual-LSTM Network; IEEE: New York, NY, USA, 2021.
  13. Talukder, D.; Jahara, F. Real-time bangla sign language detection with sentence and speech generation. In Proceedings of the 2020 23rd International Conference on Computer and Information Technology (ICCIT), Tejgaon, Dhaka, 19–21 December 2020; IEEE: New York, NY, USA, 2020; pp. 1–6.
  14. Wang, H.; Chai, X.; Hong, X.; Zhao, G.; Chen, X. Isolated sign language recognition with grassmann covariance matrices. ACM Trans. Access. Comput. (TACCESS) 2016, 8, 1–21.
  15. Camgoz, N.C.; Hadfield, S.; Koller, O.; Ney, H.; Bowden, R. Neural sign language translation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7784–7793.
  16. Venugopalan, A.; Reghunadhan, R. Applying Hybrid Deep Neural Network for the Recognition of Sign Language Words Used by the Deaf COVID-19 Patients. Arab. J. Sci. Eng. 2022, 48, 1349–1362.
  17. Kamruzzaman, M. Arabic sign language recognition and generating Arabic speech using convolutional neural network. Wirel. Commun. Mob. Comput. 2020, 2020, 3685614.
  18. Khan, N.S.; Abid, A.; Abid, K. A novel natural language processing (NLP)–based machine translation model for English to Pakistan sign language translation. Cogn. Comput. 2020, 12, 748–765.

19. Talukder, D.; Jahara, F.; Barua, S.; Haque, M.M. OkkhorNama: BdSL Image Dataset for Real Time Object Detection Algorithms. In Proceedings of the 2021 IEEE Region 10 Symposium (TENSYP), Jeju, Republic of Korea, 23–25 August 2021; IEEE: New York, NY, USA, 2021; pp. 1–6.
20. Hasan, S.N.; Hasan, M.J.; Alam, K.S. Shongket: A Comprehensive and Multipurpose Dataset for Bangla Sign Language Detection. In Proceedings of the 2021 International Conference on Electronics, Communications and Information Technology (ICECIT), Khulna, Bangladesh, 14–16 September 2021; IEEE: New York, NY, USA, 2021; pp. 1–4.
21. Wadhawan, A.; Kumar, P. Deep learning-based sign language recognition system for static signs. *Neural Comput. Appl.* 2020, 32, 7957–7968.
22. Basnin, N.; Nahar, L.; Hossain, M.S. An integrated CNN-LSTM model for Bangla lexical sign language recognition. In Proceedings of the International Conference on Trends in Computational and Cognitive Engineering, Online, 21–22 October 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 695–707.
23. Ahmed, S.; Islam, M.; Hassan, J.; Ahmed, M.U.; Ferdosi, B.J.; Saha, S.; Shopon, M. Hand sign to Bangla speech: A deep learning in vision based system for recognizing hand sign digits and generating Bangla speech. *arXiv* 2019, arXiv:1901.05613.
24. Islam, M.M.; Uddin, M.R.; AKhtar, M.N.; Alam, K.R. Recognizing multiclass Static Sign Language words for deaf and dumb people of Bangladesh based on transfer learning techniques. *Informatics Med. Unlocked* 2022, 33, 101077.
25. Shurid, S.A.; Amin, K.H.; Mirbahar, M.S.; Karmaker, D.; Mahtab, M.T.; Khan, F.T.; Alam, M.G.R.; Alam, M.A. Bangla Sign Language Recognition and Sentence Building Using Deep Learning. In Proceedings of the 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), Gold Coast, Australia, 16–18 December 2020; IEEE: New York, NY, USA, 2020; pp. 1–9.
26. Angona, T.M.; Shaon, A.S.; Niloy, K.T.R.; Karim, T.; Tasnim, Z.; Reza, S.S.; Mahbub, T.N. Automated Bangla sign language translation system for alphabets by means of MobileNet. *TELKOMNIKA (Telecommun. Comput. Electron. Control.)* 2020, 18, 1292–1301.
27. Podder, K.K.; Tabassum, S.; Khan, L.E.; Salam, K.M.A.; Maruf, R.I.; Ahmed, A. Design of a sign language transformer to enable the participation of persons with disabilities in remote healthcare systems for ensuring universal healthcare coverage. In Proceedings of the 2021 IEEE Technology & Engineering Management Conference-Europe (TEMSCON-EUR), Virtual, 17–20 May 2021; IEEE: New York, NY, USA, 2021; pp. 1–6.
28. Rahaman, M.A.; Hossain, M.P.; Rana, M.M.; Rahman, M.A.; Akter, T. A rule based system for bangla voice and text to bangla sign language interpretation. In Proceedings of the 2020 2nd

International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 19–20 December 2020; IEEE: New York, NY, USA, 2020; pp. 1–6.

29. Khan, S.A.; Joy, A.D.; Asaduzzaman, S.; Hossain, M. An efficient sign language translator device using convolutional neural network and customized ROI segmentation. In Proceedings of the 2019 2nd International Conference on Communication Engineering and Technology (ICCET), Nagoya, Japan, 12–15 April 2019; IEEE: New York, NY, USA, 2019; pp. 152–156.
30. Das, S.; Imtiaz, M.S.; Neom, N.H.; Siddique, N.; Wang, H. A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier. *Expert Syst. Appl.* **2023**, *213*, 118914.
31. Miah, A.S.M.; Shin, J.; Hasan, M.A.M.; Rahim, M.A. BenSignNet: Bengali Sign Language Alphabet Recognition Using Concatenated Segmentation and Convolutional Neural Network. *Appl. Sci.* **2022**, *12*, 3933.
32. Hassan, N. Bangla Sign Language Gesture Recognition System: Using CNN Model. *Sci. Prepr.* **2022**.
33. Akash, S.K.; Chakraborty, D.; Kaushik, M.M.; Babu, B.S.; Zishan, M.S.R. Action Recognition Based Real-time Bangla Sign Language Detection and Sentence Formation. In Proceedings of the 2023 3rd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), Dhaka, Bangladesh, 7–8 January 2023; IEEE: New York, NY, USA, 2023; pp. 311–315.
34. Tazalli, T.; Aunshu, Z.A.; Liya, S.S.; Hossain, M.; Mehjabeen, Z.; Ahmed, M.S.; Hossain, M.I. Computer vision-based Bengali sign language to text generation. In Proceedings of the 2022 IEEE 5th International Conference on Image Processing Applications and Systems (IPAS), Genova, Italy, 5–7 December 2022; IEEE: New York, NY, USA, 2022; pp. 1–6.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/114811>