BANGLADESH UNIVERSITY OF PROFESSIONALS

Mirpur Cantonment, Dhaka-1216, Bangladesh

Phone: +88-02-8000261-4, Fax: +88-02-8000443, Web: www.bup.edu.bd

Call for Papers

Bangladesh University of Professionals (BUP)
Faculty of Science and Technology (FST)
Journal of FST

Volume-4, Issue-1, July 2026

ISSN:2959-4812

DOI prefix:https://doi.org/10.64494

We are pleased to announce that FST of BUP is going to publish its yearly Journal of FST, Volume-4, Issue-I very shortly. In future, required steps will be taken soon for Indexing and Ranking. For further information, contact the editorial office at the following email address: journal.fst@bup.edu.bd

The journal welcomes various types of submission format: full research articles, short communications, book reviews, policy analysis and review articles. It encourages the submission of high-quality articles concerning the fundamentals of Electrical and Electronics Engineering, Computer Science and Engineering, Information and Communication Technology, Environmental Science and Management. Some important topics are as follows:

- Electric Power Engineering
- Electric Machinery and Power Electronics
- Electro physics & Applications
- Electric Material and Semiconductor
- High Power, High Voltage and Discharge
- Micro-Electro-Mechanical Systems (MEMS)
- Nanotechnology
- Microwave Engineering
- Radar and Satellite communications
- Optical fiber communication
- Optical and EM Wave
- Sensors and Systems
- Signal Processing
- Robotics, Automation and Control
- Application of AI in Smart Education System
- Industrial Internet of Things (IoT)
- Mobile Computing for Industry
- IoT and WSN for Smart City Applications
- Cloud Computing and Networking
- Grid and Metering Infrastructure
- Smart Transportation System
- Big Data and Machine Learning
- Natural Language Processing and Text Mining
- Data Mining for Biomedical Engineering
- Electronic Health Records and Standards
- Wearable and Body Implant Technologies
- ICT in Telemedicine
- Collaborative and Cooperative Education System
- Smart Learning System
- Cloud-IoT Platforms for Small to Large Scale Farming

BANGLADESH UNIVERSITY OF PROFESSIONALS
Mirpur Cantonment, Dhaka-1216, Bangladesh
Phone: +88-02-8000261-4, Fax: +88-02-8000443, Web: www.bup.edu.bd

- Natural Language Processing
- ICT in Education 4.0
- Cyber Physical Systems
- Cyber Physical Systems for Smart Farming
- Smart Waste Management
- Environmental Management
- Environmental Pollution and Mitigation
- Environmental Chemistry
- Environmental Engineering
- Environmental Modelling
- Environmental Economics
- Environmental Technology
- Biological Pollution in Environment
- Ecology and Biodiversity
- Earth Science
- Oceanography
- Environmental Policy and Governance
- Applications of AI and Machine Learning
- Sustainable and Green Technology
- Hydrology
- Environmental Impact Assessment and Environmental Audit System
- Water Resource Management
- Occupational Health and Safety
- Integrated coastal Zone and Floodplain Management
- Climate Change Adaptation and Mitigation
- Disaster Risk Reduction and Disaster Management
- Sustainable Urban Planning and Development
- Sustainable Energy Management
- Agriculture and Environment

All are cordially invited to submit articles for publication in the said journal. The Editorial Board will meet to review the articles for the journal soon after receiving those.

Sincerely yours

Brigadier General Sufi Md Ataur Rahman, ndc, psc
Chief Editor, Journal of FST
Dean
Faculty of Sciences and Technology
Bangladesh University of Professionals

BANGLADESH UNIVERSITY OF PROFESSIONALS
Mirpur Cantonment, Dhaka-1216, Bangladesh
Phone: +88-02-8000261-4, Fax: +88-02-8000443, Web: www.bup.edu.bd

Guidelines for Authors

Bangladesh University of Professionals (BUP)
Faculty of Science and Technology (FST)
Journal of FST
Volume-4, Issue-1, July 2026
ISSN:2959-4812
DOI prefix:https://doi.org/10.64494

Before the Beginning

Authors are requested to review the guidelines and template before their submission. All information will be available at the website of Bangladesh University of Professionals through this link www.bup.edu.bd/publications. After considering all the guidelines, authors can send their manuscripts and other required documents to the following email address: journal.fst@bup.edu.bd

Submission Checklist

You can use this list to carry out a final check of your submission before you send it to the journal for review.

Please ensure that the following items are present:

One author has been designated as the corresponding author with contact details:

- E-mail address
- Full postal address

The Manuscript should have the following information:

- Title Page (Title of the study, Corresponding author's details, Co-author's details)
- Abstract (maximum 250 words)
- Keywords (maximum 6)
- Main body (Introduction, Methodology, Result and discussion, Conclusion)
- References (APA 7 format or latest edition)
- All figures (including relevant captions)
- All tables (including titles, description, footnotes)
- Indicate clearly if color should be used for any figures in print
- Graphical Abstracts / Highlights files (where applicable)
- Page limit: Maximum 20 Pages.
- Plagiarism: Less than 20%.
- Supplemental files (where applicable)
- Please, follow the template provided on the website for this journal.

Authors should also consider the following points before the submission of their manuscript:

- This journal considers manuscript written in English.
- References follow this journal's prescribed format.
- All references mentioned in the Reference List are cited in the text.
- Permission for 3rd party information and ethical approval (where applicable).
- Manuscript should be 'spell checked' and 'grammar checked'.
- Permission should be obtained for the use of copyrighted materials from other sources (including the Internet).
- Declaration of *Conflict of Interest.*

Declaration of Submission and Verification

Submission of an article implies that the work described has not been published previously, that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronic version without the written consent of the copyright holder.

For verifying originality, your article may be checked by the originality detection service/ Similarity Checker.

Word Limit

Articles submitted to **Journal of FST** should normally have a maximum word count of 2500 for short communications, 5000 for policy analyses, 6000 for research articles and 8000 for synthesis or review articles. Flexibility with this limit may be exercised at the discretion of the handling Editor based on the quality of the Article. In general, authors are requested to write concisely and avoid overly long articles.

Essential title page information

- *Title.* Concise and informative. Titles are often used in information-retrieval systems. Avoid abbreviations and formulae where possible.
- *Author names and affiliations.* Please clearly indicate the given name(s) and family name(s) of each author and check that all names are accurately spelled. Present the authors' affiliation addresses (where the actual work was done) below the names. Provide the full postal address of each affiliation, including the country name and, if available, the e-mail address of each author.
- *Corresponding author.* Clearly indicate who will handle correspondence at all stages of refereeing and publication, as well as post-publication. This responsibility includes answering any future queries about Methodology and Materials. **Ensure that the e-mail address is given and that contact details are kept up to date by the corresponding author.**

- *Present/permanent address.* If an author has moved since the work described in the article was done, or was visiting at the time, a 'Present address' (or 'Permanent address') may be indicated as a footnote to that author's name. The address at which the author actually did the work must be retained as the main, affiliation address.

References

Please follow APA style.

Submission

The last date of manuscript submission for the current issue is **30th November 2026.** All correspondence, including notification of the Editor's decisions and requests for revision will be sent by e-mail. For further information contact the editorial office at the following email address:

journal.fst@bup.edu.bd

Final Accepted Papers

Authors will submit their final accepted articles through Word and LATEX format which will be provided later. Author must submit all figures in both color and gray scale separately.

N. B: All finally accepted article will be provided DOI.

Sincerely yours

Brigadier General Sufi Md Ataur Rahman, ndc, psc
Chief Editor, Journal of FST
Dean
Faculty of Sciences and Technology
Bangladesh University of Professionals

Bangla Sign Language Recognition: A Comprehensive Review of Machine Learning Approaches and Data Sources

Jasiya Fairiz Raisa[1]*, Rumana Yasmin[2]

Abstract

Sign language is the primary medium of communication for deaf and dumb individuals, but it is difficult to interpret for every demographic, which makes communication extremely difficult. Bangla is among the most widely spoken languages worldwide, and substantial research on Bangla Sign Language (BdSL) has emerged to address this issue. In recent years, researchers have been working to automate BdSL recognition using different techniques. This review paper evaluates research trends in BdSL by comparing the features and evaluation outcomes of various systems and approaches applied to both existing and novel datasets. We have gathered and integrated metadata from datasets encompassing all BdSL alphabets and numbers implemented to date. The analysis of this paper shows that most suggested models work well on images with static and single-handed signs, but performance drops in complicated backgrounds. Additionally, we concentrated on identifying insights and parallels within the existing systems, identifying research gaps, and suggesting potential future directions.

Keywords: Sign Language, Bangla Sign Language Recognition, Sign Language Datasets, Machine Learning, Deep Learning, Computer Vision.

1. Introduction

Sign Language is the fundamental communication method both for hearing and speech impaired individuals, as they are unable to communicate verbally or in alternative ways. This mode of communication encompasses many hand gestures, body postures, and facial expressions, each signifying a distinct meaning. Numerous sign languages exist globally, such as American Sign Language (ASL), Japanese Sign Language (JSL), British Sign Language (BSL), Austrian Sign Language (Ö GS), and Bangladeshi Sign Language (BdSL). They can all be further classified into single-handed, double-handed, static, and dynamic signs. Communication prevails as a challenge for the deaf and dumb individuals,

* Corresponding Author Email: jfraisa@bup.edu.bd

[1] Department of Information and Communication Technology, Bangladesh University of Professionals, Dhaka, Bangladesh

[2] Department of Computer Science and Engineering, Bangladesh University of Professionals, Dhaka, Bangladesh

https://doi.org/10.64494/JFST/jr/2025/03/53-86

especially in underdeveloped regions that lack the necessary support. Human-computer interface (HCI) and machine learning (ML) systems offer potential solutions by transcribing sign language into a target language using vision or sensor-based approaches. Vision-based systems capture hand gestures or signs from images or videos, while sensor-based systems detect the movement of the hands using any signs through gyroscopes and accelerometers. Bangla Sign Language (BdSL) recognition systems have shown significant progress in recent research, promoting advanced investigation for better and more effective systems. This research article reviews existing studies that developed BdSL recognition systems and relevant datasets to identify potential future research directions in this sector. Table 1 summarizes existing review articles on ML-based sign language recognition (SLR) systems. Many reviews focus on generic SLR approaches but overlook the diversity of sign languages. Given each language's unique characteristics, language-specific reviews are crucial for understanding the challenges and advancements. The literature on Bangla Sign Language (BdSL) recognition is significantly lacking, both in general reviews and those focused on individual languages. This review aims to fill that gap by providing a comprehensive analysis of BdSL recognition systems and datasets. The contributions of this research have been outlined below:

a. It offers a methodical and comparative analysis of all methodologies pertaining to BdSL recognition systems and databases.

b. The advantages and disadvantages of the existing systems have been examined and explicitly discussed in this research along with each paper's contribution.

c. Suggested research directions have been mentioned for future studies.

Section 2 outlines the generic framework of Sign Language Recognition Systems. Section 3 addresses ML techniques and relevant processing techniques for BdSL recognition. Section 4 explores different Bangla SLR datasets and benchmarks. Section 5 describes the existing Machine Learning approaches for Bangla SLR. Section 6 provides a comparative analysis of the BdSL recognition techniques and datasets and analyzes their strengths and limitations. Lastly, Section 7 concludes the paper with necessary recommendations for future research.

Table 1: Summary of the existing machine learning-based SLR systems

Reference	Area of Focus	Methodology	Key Findings
Das et al., 2021	Generic SLR systems and ML and DL in SLR	General framework for SLR, comparing techniques	SVM, KNN effective for static signs; existing challenges for dynamic signs
Elakkiya, 2021	ML techniques for SLR	Review of 240 approaches, including ML and DL techniques	ML methods show improvements in accuracy, challenges with continuous signs
Al-Qurishi et al., 2021	Deep learning methods for SLR	Review of DL models like CNN, LSTM, and hybrid methods	DL models improve dynamic sign recognition; real-time challenges remain
Tao et al., 2024	Both traditional and deep learning methods for SLR	Review datasets, techniques, and performance metrics	SVM, CNN, hybrid methods effective, limited progress with continuous signs
Renjith nd Manazhy, 2024	Multiple Language specific approaches for SLR	Systematic literature review focusing on ML techniques	SVM and Neural Networks are effective; limited focus on real-time systems
Madhiarasan and Roy, 2022	SLR systems using traditional methods and DL approaches	Comparison of SLR methods and datasets	DL approaches improve performance; existing limitations with large datasets
Subburaj and Murugavalli, 2022	Vision-based approaches for generic SLR using deep learning	Review of vision-based SLR techniques, including CNN and RNN models	Vision-based approaches improve non invasive recognition; challenges in real time systems
Adeyanju et al., 2021	ML methods for SLR, including SVM and deep learning approaches	Bibliometric analysis, comparative review of SLR methods	ML methods improve accuracy; challenges remain with signer independence
Alam et al., 2024	Machine and deep learning for sign language detection on smartphones	Systematic review of smartphone-based SLR techniques	CNNs show promise for smartphone-based SLR; limited datasets
Moustafa et al., 2024	Arabic Sign Language (ArSL) recognition systems	Systematic review of ArSL recognition approaches	Vision-based systems, especially Kinect, are common; challenges with dynamic signs
Zahid et al., 2022	Urdu SLR using machine learning classifiers	Systematic review of machine learning techniques applied to Urdu Sign Language	SVM and Neural Networks are most used; dataset limitations
Khatun et al., 2021	Focus on BdSLR	Systematic review of BdSLR research from 2002-2021	BdSLR systems perform well for static signs, dynamic sign challenges

2. Sign Language Recognition Framework

Sign Language Recognition (SLR) system is designed to ensure better communication between the hearing-impaired and the general population (Rastgoo et al., 2021). To develop an effective SLR system, a structured framework is necessary with ML integration. This section outlines the generic and ML-based architecture for SLR systems.

2.1 General Architecture of SLR Systems

SLR architecture is comprised of multiple interconnected components. Each one is essential for processing and interpreting sign language motions. SLR systems use ML models to recognize and classify signs accurately. A standard SLR system consists of the following key components:

a. Input Module: Records signs using different methods like video cameras, depth sensors, or wearable motion sensing devices.

b. Preprocessing Module: The collected data is cleaned and normalized for improving its reliability and quality.

c. Feature Extraction Module: Extracts important features like hand shape, motion trajectory, and facial expressions from the processed data.

d. Classification Module: Implements different classification methods to classify the extracted features into relevant sign language labels.

e. Output Module: Renders recognized signs into readable format like text or speech for effective communication.

These components combinedly develop the structure of a generic SLR system that ensures seamless recognition and transcription of sign language gestures.

2.2 Machine Learning-Based Sign Language Recognition System Architecture

The architecture of ML-based SLR system is quite similar to the generic framework, however, it follows a data-driven pipeline.

a. Data Acquisition: The first step in the ML-based SLR system requires capturing the gestures through different input modalities:

- Video-Based Input: High-resolution cameras capture hand movements, facial expressions, body posture that are compatible with ML models.
- Sensor-Based Input: Wearable sensors like accelerometers and electromyographic (EMG) sensors can detect hand and finger movements in real-time.

- Hybrid Approaches: Video and sensor data can be combined to enhance sign language detection accuracy with the help of visual and motion-based information.

b. Preprocessing: Preprocessing handles the noise and inconsistencies in the raw data and enhances data quality:

- Noise Reduction: Reduction of background noise and irrelevant factors.
- Frame Normalization: Standardization of video frames for uniformity.
- Hand and Face Detection: Identifying and isolating key areas (hands and face).
- Temporal Alignment: Ensuring that sequential frames are properly arranged for capturing dynamic movements.

c. Feature Extraction: Essential features of sign gestures are extracted for classification:

- Hand Shape and Position: Recognizing hand are oriented.
- Motion Trajectory: Monitoring hand movement to understand dynamic signs.
- Facial Expressions: Recognizing emotive expressions that comprise of sign gesture meaning.
- Depth Information: Depth cameras are implemented to record 3D positions for easier spatial understanding.

d. Classification: ML models classify the extracted features into corresponding sign language labels:

- Traditional Methods: Support Vector Machines (SVM), Hidden Markov Models (HMM), and k-Nearest Neighbors (k-NN).
- Deep Learning (DL): CNNs, and Recurrent models like RNNs and Long Short-Term Memory (LSTM) models are used for detecting spatial features and temporal dependencies respectively. Transformer-based models provide advanced sequential modeling.

e. Prediction and Output Generation: After classification, the system converts model predictions into human-interpretable formats, such as:

- Textual Representation: Displaying recognized signs as text.
- Spoken Output: Converting signs into speech through text-to-speech systems.

Modern SLR systems are highly versatile and accurate owing to advanced ML techniques. This allows these systems to be applicable for real-world implications. Fig 1 visualizes the framework of the ML-based SLR system.

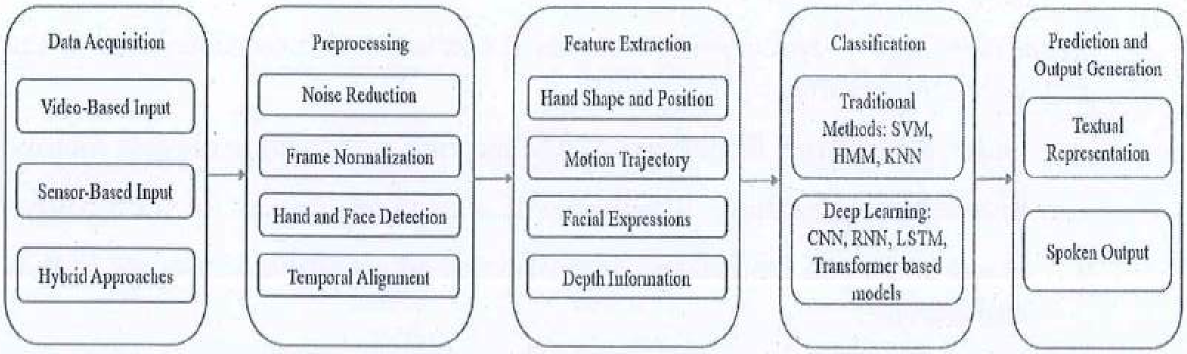

Figure 1: Workflow of machine learning-based SLR system.

3. Classification of Bangla Sign Language Recognition Using Machine Learning

ML-based BdSL classification has advanced drastically recently (Khatun et al., 2021). From traditional to DL, different ML models have been implemented to increase SLR system accuracy. In this section, notable ML algorithms are examined with advantages and disadvantages and their implementability in BdSL recognition.

3.1 Overview of ML Techniques for Bangla Sign Language Recognition

The ML-based BdSL recognition systems can be divided into two parts: traditional ML and DL models.

3.1.1 Traditional ML Techniques for BdSL Recognition

Many traditional ML algorithms are found in the early BdSL works. These models include:

a. Support Vector Machines (SVM): SVM is vastly used for classifying BdSL signs like hand trajectory, shape, and motion (Renjith and Manazhy, 2024). Small datasets provide better classification results with SVM. However, this model struggles with real-time recognition of continuous sign gestures. SVM's ability to efficiently classify small, high-dimensional datasets makes it suitable for recognizing static hand gestures in BdSL.

b. k-Nearest Neighbors (k-NN): k-NN is a simple and efficient model, that has been found useful for categorizing isolated BdSL signs, based on feature similarity (Raihan et al., 2024). Real-time application development with this model is difficult due to its processing cost rising with larger datasets. k-NN is well-suited for BdSL due to its simplicity in classifying gestures based on feature similarity, especially for isolated signs.

c. Hidden Markov Models (HMM): The temporal dependencies in dynamic BdSL signs have been modeled using HMM (Sandjaja et al., 2024). Nonetheless, HMM is less successful than DL- based sequence models for its restricted learning capacity. HMM is effective in modeling the sequential nature of BdSL signs, capturing temporal dependencies in dynamic gestures.

d. Random Forest (RF): RF is used for feature selection and classification. Though it produces reliable results, complicated spatiotemporal sign gestures are handled less effectively (Das et al., 2023). RF's ensemble learning approach is beneficial for classifying diverse BdSL features but struggles with complex spatiotemporal gestures.

e. Extreme Gradient Boosting (XGBoost): For BdSL classification, ensemble method found to improve prediction accuracy over conventional tree-based models. But this model requires significant feature engineering (R. Islam et al., 2023). XGBoost's strong performance in classification tasks makes it ideal for BdSL, provided substantial feature engineering is applied.

3.1.2 Deep Learning Models for BdSL Recognition

The performance of BdSL recognition has greatly improved with recent developments in deep learning. Several popular deep learning models are as follows:

a. Convolutional Neural Networks (CNNs): Static BdSL signs are commonly recognized from pictures or video frames using architectures such as VGGNet, ResNet, and MobileNet (Khatun et al., 2021). While CNNs excel at extracting spatial features, they are unable to capture temporal relationships in dynamic sign. CNNs excel at recognizing static BdSL signs from image frames by extracting spatial features, though they cannot model dynamic gestures effectively.

b. Recurrent Neural Networks (RNNs) and Long-Short-Term Memory (LSTM): RNNs and LSTM networks are used to recognize dynamic BdSL gestures by processing sequential video frames (Mhatre et al., 2022). These models effectively learn temporal dependencies but suffer from long training times and the vanishing gradient problem. RNNs and LSTM are well-suited for recognizing dynamic BdSL gestures by learning temporal relationships across sequential frames.

c. Transformer-Based Models: Vision Transformers (ViTs) and attention-based architecture have recently emerged as powerful tools for sign language recognition. They outperform traditional sequence models in handling long-term dependencies. Nevertheless, they require massive computational resources and significant datasets. These models can capture long-range dependencies in BdSL signs, offering high accuracy for complex gestures.

3.1.3 Evaluation of ML Models for BdSLR Systems

This section provides a comparative evaluation of the prior discussed ML and DL models for BDSLR Systems. Table 2 compares the key features, advantages, and challenges of each model for BdSLR. Traditional machine learning models are efficient for small datasets and static signs but struggle with real-time and dynamic sign recognition due to limited temporal dependency handling. While less computationally demanding, their scalability requires further refinement. Deep learning models like CNN, RNN, LSTM, and Transformers show potential but may need optimization for real-time deployment, particularly in resource-constrained environments. This evaluation highlights the trade-offs between model performance and computational efficiency for static and dynamic signs.

Table 2: Functional and Performance Comparison of ML Models for BdSLR Systems Based on Key Attributes

Model	Real-time Recognition	Temporal Dependencies	Static Signs Detection	Dynamic Signs Detection	Computational Efficiency	Scalability & Transparency
SVM	N	N	Y	N	Y	N
k-NN	N	N	Y	N	N	N
HMM	N	Y	N	Y	N	Y
RF	N	N	Y	N	Y	Y
XGBoost	N	N	Y	Y	N	Y
CNN	Y	N	Y	N	Y	Y (with optimizations)
RNN	Y	Y	N	Y	N	Y (with optimizations)
LSTM	Y	Y	N	Y	N	Y (with optimizations)
Transformer	N	Y	N	Y	N	Y (with optimizations)

Although ML-based BdSL identification improves communication for the deaf people who speak Bangla, it has drawbacks such as a lack of datasets, high processing costs, and accessibility problems. Inclusion can be enhanced by expanding datasets, refining models, and reducing costs.

4. Bangla SLR Datasets and Benchmarks

Sign language datasets are pivotal for ML-based systems, as they ensure the application and validity of algorithms. Data acquisition has been simplified with the AI model development. This section reviews existing publicly available BdSL datasets that are used for ML-based BdSL systems.

4.1 Features and Characteristics of Bangla Sign Language Datasets

A significant number of datasets have been developed to support BdSL research, covering different sign levels and modalities. We have reviewed 16 datasets in total, that include image and video data and focus on both static and dynamic signs. Most of the reviewed datasets are developed in controlled environments, while few others incorporated real-world diversity. Based on sign granularity, the existing Bangla Sign Language (BdSL) datasets can be broadly categorized into three groups: character-level, word-level, and sentence-level datasets.

4.1.1 Character-level Datasets

Character-level datasets mainly comprise isolated static images of Bangla alphabets, digits, or special characters, often with class labels and sometimes bounding boxes.

Ishara-lipi (M. S. Islam et al., 2018) includes 1,800 images of 36 Bangla sign characters (vowels and consonants) taken in controlled lighting. The dataset consists of clear class labels, but no dynamic gestures, real-world testing, or advanced annotations. These characteristics limit its use for complex SLR systems.

Ishara Bochon (M. S. Islam et al., 2018) specifically focuses on Bangla numerals, with 1000 grayscale images (100 images for 10 digits each). It is incompetent to be used for dynamic and scalable systems due to its small size and focus on static numerals.

Rafi et al., 2019 developed a Bangla Sign alphabet for 9 vowels and 27 consonants, containing 12581 images evenly distributed across the 38 classes. While the dataset is balanced and well-processed, it does not have significant diversity of backgrounds and recording context.

Hoque et al., 2021 initially worked with 2,712 images, later expanding to over 473,000 images in BdSL36v2. It features 36 Bangla sign letters with balanced class distribution and includes background augmentation, varied lighting, angles, and perspectives. However, some classes have high similarity, which impacts accuracy in real-world scenarios.

Tasmere et al., 2020 examined the Bangla Sign Digits Dataset for real-time digit recognition (0–9), containing 1,674 images with uneven class distribution collected via webcam under varied conditions. While it offers high accuracy and effective preprocessing, its small size and lack of alphabet or sentence-level data limit its broader application in BdSL.

Shongket (S. N. Hasan et al., 2021) offers 5,820 images of Bangla digits (10) and letters (36), with uniform distribution for digits and varying distributions for letters. Captured under diverse lighting and hand orientations, the dataset offers a variety of environmental factors for static gestures. However, it lacks dynamic or continuous signing, limiting its use for real-time applications.

Talukder et al., 2021 presented Okkhornama, a dataset with 12,000 high-resolution images of 46 Bangla signs (10 digits, 6 vowels, 30 consonants), with about 260 images per class, captured under varying conditions. It includes bounding box annotations in YOLO format, suitable for real-time object detection. However, it is limited to static single-hand gestures and lacks dynamic or sentence-level data.

Jim et al., 2023 developed KU-BdSL, an open dataset with 1,500 images of 38 Bengali consonants (30 hand signs) captured under varied lighting and backgrounds, with DarkNet-format annotations and bounding boxes for deep learning-based detection. While it features high signer diversity, it excludes vowels, words, and dynamic gestures, limiting its use for continuous sign language recognition.

The BdSL 49 dataset (Hasib et al., 2023) includes 29,490 images across 49 classes (37 alphabets, 10 digits, and 2 special characters), with a uniform class distribution and annotations for both detection and recognition tasks. Captured with various smartphone cameras, it offers real-world diversity in lighting and backgrounds. However, it only includes static single-hand gestures and has a limited participant pool, which limits its generalizability.

Rayeed et al., 2023 developed BdSL47, a depth-based dataset comprising 47,000 images and 470 CSV files with 3D hand key points for 37 Bangla alphabets and 10 digits. It offers high diversity in hand shapes, lighting, and gesture orientations, supporting multimodal deep-learning research. However, its static gestures and controlled settings limit its applicability in dynamic, real-world scenarios.

Hadiuzzaman et al., 2024 developed BAUST Lipi, a dataset of 18,000 images (600 per character) for 36 Bangla characters, with a balanced distribution and diverse lighting and backgrounds. While large and well-preprocessed for machine learning, it focuses on isolated gestures and lacks dynamic or sentence-level data, limiting its real-world applicability.

4.1.2 Word-level Datasets

Word-level datasets include isolated and continuous video sequences of common words or glosses, capturing more complex gestures that go beyond individual characters toward natural language use. Moving to more recent datasets, M. M. Islam et al., 2022 developed BDSLW-11, which contains 1,105 images of 11 common Bangla sign words. While small, it achieves high classification accuracy and has been preprocessed for optimal model performance. However, its limited size and lack of background and lighting diversity may restrict its use in real-world applications.

M. S. Islam et al., 2023 introduced A15, a multi-view BdSL dataset containing 9,204 videos (6,000 for 350 words and 3,204 for 107 glosses), captured in diverse real-life and artificial settings, with pose and landmark annotations to enhance model robustness. While the multi-view data improves generalization, limited real-world testing and high computational costs present challenges for real-time applications.

Sign-BD Sams et al., 2023 contains 6,000 video clips of 200 common Bangla sign words, annotated with 2D body pose key points. While suitable for sign classification, translation, and synthesis, its limited vocabulary, participant pool, and lack of continuous sentence-level signing restrict its broader applicability.

BdSLW60 by Rubaiyeat et al., 2025 is another video-based dataset of 60 Bangla sign words with 9,307 trials and 2D pose key points from OpenPose. While suitable for sign translation tasks using GANs, its limited vocabulary of 200 words and small participant pool restricts its representation of the broader Bangla sign language community.

4.1.3 Sentence-level Datasets

Sentence-level datasets consist of continuous signing in full sentences or longer sequences, with detailed temporal annotations. These datasets reflect natural sign language flow and are key for real-time communication applications.

BTVSL (Zeeon et al., 2024) is a large-scale dataset for sentence-level recognition featuring 60 hours of video, 24,085 annotated sentences, and 340,172 words. Extracted from Bangla news, it offers detailed annotations and 48,623 unique tokens, aiding sign language translation. However, challenges in linguistic diversity and gloss annotations limit its applicability.

Table 3 compares the progression of BdSL datasets, highlighting their development for ML-based BdSL recognition systems.

Table 3: Comparison of Different BdSL Datasets

Dataset	Sign Level	No. of Samples	Sign Sequence Type	Data Modality and Type	Annotation
Ishara-lipi (M.S. Islam et al., 2018)	Charactr (36 Bangla sign letters)	1800	Isolated	Image, RGB	Numerically labeled images (no bounding boxes/metada a)
Ishara Bochon (M.S. Islam et al., 2019)	Charactr (10 sign Bangla digits)	1000	Isolated	Image, Grayscale	Class labels in folders (no metadata)
Bangla Alphabet Dataset (Rafi et al., 2019)	Character (38 Bangla Bangla sign letters)	12,581	Isolated	Image, RGB	Class identifiers for characters (no bounding boxes)
BDSL36 (Hoque al., 2021) et	Charactr (36 Bangla sign letters)	2712 (initial), 473,662 (BdSL36)	Isolated	Image, RGB	Classlabels bounding boxes with
Bangla Sign Digits (Tasmere et al., 2020)	Character (10 Bangla sign digits)	1674 total images	Isolated	Image, RGB	Class labels for digits (organized by class)
Shongket (S. N. Hasan et al., 2021)	Character (36 Bangla sign letters, 10 sign Bangla Digits)	5820 (4320 letters, 1500 digits)	Isolated	Image, RGB	Indexed class labels (no bounding boxes)
Okkhornama (Talukder et al., 2021)	Character (46 Bangla signs: 10 digits, 36 letters)	12,000	Isolated	Image, RGB	YOLO format bounding boxes with class labels
KU-BdSL (Jim et al., 2023)	Character (38 sign Bengali letters)	1500	Isolated	Image, RGB	DarkNet format bounding boxes
BdSL 49 (Hasib et al., 2023)	Character (37 alphabets, 10 digits,2 special characters)	29,490	Isolated	Image, RGB	Bounding boxes for detection; class folders for recognition
BdSL47 (Rayeed et al., 2023)	Character (37 alphabets, 10 digits)	47,000 images; 470 CSV files with 3D keypoints	Isolated	Image, RGB with 3D depth coordinates	Hand variations from 10 participants

BAUST Lipi (Hadiuzzaman et al., 2024)	Character (36 Bangla alphabets: 30 consonants, 6 vowels)	18,000	Isolated	Image, RGB	Numerically labeled images (no bounding boxes/metadata)
BDSLW-11 (M. M. Islam et al., 2022)	Word (11 common Bangla sign words)	1105	Isolated	Image, RGB	Labeled dataset with class names for words
A15 (M. S. Islam et al., 2023)	Words: 350 isolated words; Sentences: 107 sign glosses	9204	Isolated word level and continuous gloss- level videos	Video, RGB	Class labels for videos, pose/landmark annotations
Sign-BD (Sams et al., 2023)	Word (200 common Bangla sign words)	6000 video clips	Continuous	Video, RGB + 2D body pose keypoints	Gloss labels with pose/speech annotations
BdSLW60 (Rubaiyeat et al., 2025)	Word Bangla words) (60 sign	9307 video word-level signs of	Continuous	Video, RGB	Frame numbers, hand dominance, class labels (JSON/text)
BTVSL (Zeeon et al., 2024)	Sentence (24085 sentences)	60 hours of video	Continuous	Video, RGB	Sentence-level annotations (text and timestamps)

Most BdSL datasets focus on isolated character-level signs, likely because they are easier to collect and annotate. Word-level datasets, though fewer, are growing and more complex due to video and gloss annotations. Sentence-level datasets are rare and challenging but essential for real-time, fluent sign language recognition. BdSL datasets are essential for recognition systems but remain limited by class imbalance and a focus on static gestures, restricting dynamic applicability. Future datasets should prioritize dynamic gesture recognition, multimodal data, real-time performance, signer diversity, and cross-dataset compatibility to enhance robustness and practical use.

5. Existing Machine Learning Approaches for Bangla SLR

A brief discussion of the existing studies on Bangla Sign Language recognition is provided in this section.

5.1 Traditional Machine Learning Approaches for Bangla SLR

This section reviews traditional machine learning-based approaches for Bangla SLR, including PCA, SVM, and k-NN, highlighting their applications in static gesture classification and computational efficiency.

A Computer Vision-based BdSL Recognition System is proposed by Begum and Hasanuzzaman, 2009 to bridge communication gaps between deaf/mute individuals and non-sign language users. The system employs PCA to recognize 6 Bengali vowels and 10 Bengali numerals from hand gestures. A key challenge addressed is that unlike many other sign languages, BdSL often uses both hands (except for numbers), increasing recognition complexity. The training dataset consists of 240 vowel images and 400 number images. Despite the complexity of two-handed gestures, the system achieves satisfactory recognition accuracy, demonstrating the feasibility of PCA-based sign language recognition.

Rahman, 2012 developed a BdSL alphabet recognition system that classifies 36 static hand gestures with 80.90% accuracy. It employs a feed-forward ANN using a sigmoid activation function and mean square error as the loss function.

Uddin et al., 2017 presented a model for BdSL character recognition that is based on image processing. The bag of words approach, which includes visual vocabulary learning, feature quantization, and image representation, was used to extract features from RGB images after they had been converted to YCbCr components. These attributes were then analyzed using a SVM multiclass classifier, which produced an accuracy of 86

Shanta et al., 2018 presented a SIFT and CNN-based Bangla Sign Language identification system. The image was first segmented using skin masking, and then Canny edge detection was used. SIFT extracted scale-invariant features, which were clustered using k-means before being processed by a CNN for classification.

A BdSL digits recognition method was proposed by M. M. Hasan and Ahsan, 2019 using HOG features and a multi-class ONE VS ONE SVM. The procedure achieved 94.74% accuracy with decreased time for sign recognition. A model for categorizing the BdSL alphabet was presented by Haque et al., 2019. They obtained a recognition accuracy of 77.88% using a dataset of 130 samples that they constructed themselves. PCA was used to reduce the dimensionality of the data and extract features.

A single-handed Bengali SLR system based on HOG features was proposed by Tabassum et al., 2020. They created a collection of 1400 photos and used lightness smoothing and histogram equalization to improve the quality. A K-Nearest Neighbor classifier was trained using grayscale images and HOG features, which yielded an accuracy of 91.1%.

Youme et al., 2021 emphasized generalization challenges in BdSL recognition and dataset limitations. They used the Ishara-Lipi Dataset (35 classes, 978 images) and the BdSL Dataset (35 classes, 23,786 images). The VGG19 model with SphereFace loss achieved 55.93% and 47.81% inter-dataset accuracy, highlighting the need for larger, more diverse datasets and better generalized models.

Table 4 summarizes the different Bangla SLR systems implemented using traditional ML based approaches, along with the methodology used and model accuracy.

Table 4: Summary and comparison of different BdSL recognition systems using traditional ML.

Reference	Dataset	Total Data	Methodology/Algorithm Used	Accuracy
Begum and Hasanuzzaman, 2009	Custom	1,920	PCA + Euclidean Distance-Based Classification	N/A
Rahman, 2012	Custom	N/A	Feed-forward ANN with sigmoid activation and mean square error loss	80.90%
Uddin et al., 2017	Custom	N/A	Bag of Words extraction + SVM multiclass classifier	80%
Shanta et al., 2018	Custom	1,700	SIFT feature extraction + k- means clustering + CNN	Not specified
M. M. Hasan and Ahsan, 2019	Ishara-Bochon	1,000	HOG feature extraction + ONE VS ONE multi-class SVM	94.74%
Haque et al., 2019	Custom	130	PCA for feature extraction + K-Nearest Neighbor (KNN) classifier	77.88%
Tabassum et al., 2020	Custom	1,400	HOG feature extraction + k- NN	91.10%
Youme et al., 2021	Ishara-Lipi, BdSL	978 (Ishara-Lipi), 23,786 (BdSL)	VGG19 with SphereFace loss; angular margin-based loss functions	55.93% (inter-dataset)

5.2 Deep Learning Approaches for Bangla SLR

In this section, we explore deep learning-based techniques for Bangla SLR, analyzing how modern neural networks address the limitations of traditional approaches in recognizing dynamic and multi-handed signs. Table 5 summarizes the different Bangla SLR systems implemented using Deep Learning based approaches, along with the methodology used and model accuracy.

To facilitate both text-to-sign and sign-to-text conversion, Rony et al., 2018 presented a dual-purpose translator for one-handed Bangla Sign Language. They achieved 92.85% accuracy using InceptionV3 as the foundational CNN model with transfer learning.

For real-time usability, Hossen et al., 2018 developed a CNN-based BdSL recognition algorithm employing a broad collection of natural-color images with varying lighting and backdrops. On a small collection of 1147 images that cover 37 signs, this study achieved 84.68% accuracy employing a pre-trained VGG16 model.

Khan et al., 2019 developed a real-time, low-cost Sign Language Translator on a Raspberry Pi 3B, detecting five signs from webcam videos. A self-built dataset was augmented and pre-processed before training a CNN classifier. The system achieved 94% accuracy in sign classification. A large and diverged dataset, along with a deep learning-based recognition model, was proposed in Rafi et al., 2019. They considered 38 BdSL alphabets, applied a pre-trained VGG19 transfer learning model for classification, and achieved 89.60% test accuracy. A computer vision-based system employing Deep CNN was presented by Ahmed et al., 2019. It can recognize Bangla Sign digits and convert them into Bangla speech with an accuracy of up to 92%.

Shamrat et al., 2021 created a CNN-based Bangla Numerical SLR system to assist hearing-impaired people in communicating numerical signs through webcam-captured images. A dataset of 310 images (10 classes, 31 images each) was created, and the CNN model achieved 99.8% accuracy.

The study performed by Surjo et al., 2023 compares CNN effectiveness in recognizing single-handed and dual-handed BdSL. The KU-BdSL dataset (30 alphabets) was used for single-handed BdSL, while the BdSL-49 dataset (36 alphabets) was used for dual-handed BdSL. Among three pre-trained CNN models, VGG16 achieved the highest accuracy: 98% for single-handed and 90% for dual-handed BdSL. However, the study is limited to alphabet recognition, excluding vocabulary and sentence-level gestures.

An automatic BdSL detection system was proposed by Siddique et al., 2023 using deep learning and the Jetson Nano edge device. A custom dataset of 49 classes (3,760 images) with varying backgrounds and lighting was created, along with the Okkhornama dataset for training. Three models—Detectron2, EfficientDet-D0, and YOLOv7—were implemented. Detectron2 achieved the highest accuracy with a mAP@.5 of 94.915 and an AP of 54.814, while the mAP@.5 of YOLOv7 ranged from 85% to 97%. The YOLOv7 Tiny model was deployed on Jetson Nano for real-time detection, offering low training time and a high frame rate.

The study conducted by Abedinc et al., 2023 aimed to improve BdSL recognition by addressing challenges such as the similarity of some alphabets and the importance of hand posture. It proposed the Concatenated BdSL Network, which combines visual features from a CNN-based image network with hand key point positions from a pose estimation network to enhance recognition accuracy. The methodology achieved a test accuracy of 91.51%, which demonstrated the efficacy of including hand pose data.

An ML-based Bengali Sign Language (BdSL) interpretation system was developed by Raihan et al., 2024 as part of an approach to close the communication gap between the hearing community and deaf or non-verbal people. The model uses CNN in combination with a smartphone app for accessibility and a Squeeze Excitation (SE) block to improve recognition accuracy. The model interprets its decisions using SHapley Additive eXplanation (SHAP). The accuracy of the CNN-SE model on the KU-BdSL dataset was 99.86%.

5.3 Hybrid Model Approaches for Bangla SLR

This section discusses hybrid learning approaches for Bangla SLR, examining how the fusion of ML-based feature extraction with deep learning models creates more efficient and adaptive systems for real-world sign language interpretation applications.

A real-time BdSL digit classification technique utilizing video data was presented by Tasmere et al., 2020. The Gaussian Mixture Model (GMM) was used to recognize moving objects in training images that were taken from video frames. Prior to being fed into a deep CNN, pre-processing involved blurring, thresholding, and normalization, yielding an accuracy of 97.63%.

Using a CNN-LSTM model, Basnin et al., 2021 created a BdSL recognition system for lexical signs that can manage differences in hand size, shape, orientation, and motion. 13,400 images from 36 classifications made up the dataset. The CNN-LSTM model outperformed CNN, VGG16, VGG9, and MobileNet in terms of generalization and overfitting, achieving 90% training and 88.5% testing accuracy.

Podder et al., 2022 developed a real-time BdSL interpreter for alphabets and numerals to aid communication for hearing/speech-impaired individuals in Bangladesh. They introduced BdSL-D1500, the largest publicly available dataset with 87 classes, covering one-hand and two-hand BdSL alphabets, digits (0–9), numerals (00, 000), and "Counting." Transfer learning was applied to ResNet18, MobileNetV2, and EfficientNetB1 for BdSL recognition, alongside a semantic segmentation-based hand detection approach using M-UNet, DenseNet201 FPN, and UNet. ResNet18 achieved 99.99% accuracy and 100% specificity, while DenseNet201 FPN reached 98.644% accuracy for hand detection. Models trained with background images outperformed those using segmented hand images.

Table 5: Summary and comparison of different BdSL recognition systems using Deep Learning.

Reference	Dataset	Total Data	Methodology/Algorithm Used	Accuracy
Raihan et al., 2024	KU-BdSL	1,500	CNN with Squeeze Excitation (SE) block, SHAP analysis, smartphone app	99.86%
Siddique et al., 2023	Okkhornam, Custom	3,760	Detectron2, Efficient Det-D0, YOLOv7, Jetson Nano for real-time detection	94.915% (mAP@.5)
Abedine et al., 2023	Custom	12,581	CNN + Pose Estimation Network (Concatenated BdSL Network)	91.51%
Surjo et al., 2023	KU-BdSL, BdSL 49	1,500 (KUBdSL), 2,500 (BdSL 49)	Pre-trained CNN models (VGG16, VGG19, MobileNet)	98% (single-handed), 90% (dual handed)
Shamrat et al., 2021	Custom	310	CNN-based approach, webcam for real-time detection	99.80%
Ahmed et al., 2019	Custom	3,200	Deep CNN	92%
Khan et al., 2019	Bangla sign languagediction ary (Committee, 1994)	500 training samples (augmented to 4,000)	CNN (3 convolutional layers, 3 normalization layers, 2 fully connected layers) + Raspberry Pi	94%
Rafi et al., 2019	Custom	12,581	Transfer learning with VGG19	89.60%
Hossen et al., 2018	Custom	1,147	Transfer learning with VGG16 pre-trained on ImageNet	84.68%
Rony et al., 2018	Society of the Deaf & Sign Language Users (SDSL)	2,050	Transfer learning with InceptionV3 as backbone CNN	92.85%

For BdSL recognition, Das et al., 2023 have developed a hybrid model that combines deep transfer learning with a Random Forest classifier and pre-trained models such as VGG16, VGG19, InceptionV3, Xception, and ResNet50. Performance is further improved by grid search optimization and background removal. The model achieved 97.33% accuracy for digits and 91.67% accuracy for characters when tested on two datasets: Ishara-Bochon (numerals) and Ishara-Lipi (alphabets).

The KUNet AI-based solution was created by Jim et al., 2024 for interpreting Bengali Sign Language (BdSL) to help non-verbal and hearing-impaired people. The suggested model makes use of CNN that has been tuned by Genetic Algorithms (GA) and interprets its decisions using Explainable AI (XAI). They made use of a dataset that included 1,500 pictures of 38 Bengali consonants (30 classes, 50 pictures each). KUNet's accuracy on the KU-BdSL dataset was 99.11%. It performed better in terms of accuracy, recall, precision, and F1 score than well-known deep learning models such as AlexNet, ResNet50, VGG16, and VGG19. The model is not appropriate for real-time scenarios involving multiple gestures because it can only predict one hand gesture from an image.

Table 6 compares datasets of BdSL recognition systems from the newest to the oldest, with an emphasis on the development of hybrid model based BdSL recognition systems.

6. Discussion

This section delves into Bangla SLR (BdSLR) systems and datasets in extensive detail. Reviewed BdSLR systems are compared, with an emphasis on the approaches, strengths and limitations. Subsequently, BdSL datasets are analyzed, focusing on their diversity and challenges. The section concludes with an outline of the key observations of BdSLR systems and datasets with a trend analysis of the current research in this sector.

Table 6: Summary and comparison of different BdSL recognition systems using Hybrid models.

Reference	Dataset	Total Data	Methodology/Algorithm Used	Accuracy
Jim et al., 2024	KU-BdSL	1,500	CNN optimized by GA, XAI	99.11%
Das et al., 2023	Ishara-Bochon, Ishara-Lipi	1,075 (Ishara-Bochon), 1,005 (Ishara-Lipi)	Transfer learning (VGG16, VGG19, InceptionV3, Xception, ResNet50) + Random Forest Classifier	91.67% (characters), 97.33% (digits)
Podder et al., 2022	Ishara-Lipi, BdSL	978 (Ishara-Lipi), 23,786 (BdSL)	ResNet18, transfer learning, semantic segmentation	99.99%
Basnin et al., 2021	BdSL	13,400	CNN-LSTM with background subtraction, grayscale conversion, median filtering	88.5%
Tasmere et al., 2020	Custom	1,674	Gaussian Mixture Model (GMM) for moving object detection + Deep CNN	97.63%

6.1 Comparative Analysis of Bangla SLR Datasets

To create effective BdSL recognition systems, robust datasets that cover character, word, and sentence-level recognition are needed. This section compares the Bangla SLR datasets, noting their strengths and weaknesses in development.

6.1.1 Strengths of BdSL Datasets

The strengths of the existing BdSL datasets include large datasets, augmentation, and dynamic gesture recognition while ensuring real-world variability and high recognition accuracy. The strengths are discussed in detail as follows:

- Large and Diverse Datasets: Datasets like Hoque et al., 2021 and BdSL 49 Hasib et al., 2023 are among the largest and most diverse. They offer extensive coverage of various sign classes, backgrounds, and lighting conditions, which help in building generalized models for robust recognition.
- Comprehensive Annotations and Augmentation: Datasets like A15 by M. S. Islam et al., 2023 offer detailed annotations, including pose and landmark data, crucial for advanced tasks like continuous sign recognition. Implementing augmentation methods and techniques ensures their effectiveness for real-world applications.
- Real-World Variability: Many recent datasets like BTVSL Zeeon et al., 2024 have started to incorporate real-world contexts, like natural variations in sign gestures. This variability enhances the generalization of the datasets for practical implementation.
- Dynamic Gesture Representation: Dynamic and real-time sign gestures are essential for continuous BdSLR tasks, which are demonstrated in many datasets like Sign-BD by Sams et al., 2023.
- High Accuracy: Multiple datasets like BAUST Lipi by Hadiuzzaman et al., 2024 and Shongket by S. N. Hasan et al., 2021, exhibited impressive benchmark performance for sign language recognition, making them better suited for high-performance SLR systems.

6.1.2 Limitations of BdSL Datasets

While BdSL Datasets contain a large data volume, significant challenges remain to be resolved to ensure improved SLR performance. The key challenges for dataset development are as follows:

- Limited Dynamic Gesture Representation: Datasets like Ishara-lipi by M. S. Islam et al., 2018 and Ishara Bochon by M. S. Islam et al., 2019 focus on isolated gestures, limiting their suitability for dynamic, real-time continuous sign language applications.

- Limited Real-World Testing and Variability: A15 (M. S. Islam et al., 2023) and BdSL47 (Rayeed et al., 2023) lack sufficient real-world variability, limiting their generalizability. Moreover, most datasets are captured in controlled environments, restricting their ability to account for variations in signing speed, hand posture, and background noise.
- Participant Diversity: Many datasets, including BAUST Lipi (Hadiuzzaman et al., 2024), Ishara-lipi (M. S. Islam et al., 2018), and Okkhornama (Talukder et al., 2021), have a limited diversity of signers. This lack of demographic diversity, particularly in terms of age, gender, and ethnic background, affects the generalization capability of BdSL recognition models.
- Small Dataset Sizes: Some datasets, such as Ishara-lipi, contain a small number of samples, limiting their utility in training robust deep-learning models.
- Background and Environmental Variability: Several datasets, like Ishara Bochon, focus on controlled environments, which do not adequately represent real-world conditions.
- Missing Temporal and Contextual Data: Most BdSL datasets lack necessary temporal and contextual annotations, leading to less accurate interpretation of continuous sign language gestures.
- Lack of Multimodal Data Integration: The majority of BdSL datasets typically exclude multimodal data (e.g., motion sensors, depth cameras, or audio), reducing gesture recognition performance in complex environments. Recent datasets have started to address this issue by integrating multimodal data like depth information of gesture images (Rayeed et al., 2023).
- Linguistic Diversity: Regional variation is not yet represented in the existing BdSL datasets, which limits dataset robustness and dynamic applicability (Zeeon et al., 2024).

Table 7 highlights the significant features of the studied datasets.

To advance the BdSL research forward and develop more inclusive, precise, and real-life implementable systems, it is important to address the above-discussed limitations.

Table 7: Comparison of different Bangla Sign Language Dataset features

Dataset	Static Gestures	Dynamic Gestures	Pose/ Landmark Data	Real-World Variability	Multimodal Data
Ishara-lipi (M. S. Islam et al., 2018)	Y	N	N	Y	N
Ishara Bochon (M. S. Islam et al., 2019)	Y	Y	N	Y	N
Bangla Alphabet Dataset (Rafi et al., 2019)	Y	Y	Y	Y	Y
BDSL36 (Hoque et al., 2021)	Y	N	N	Y	N
Bangla Sign Digits (Tasmere et al., 2020)	Y	Y	Y	Y	Y
Shongket (S. N. Hasan et al., 2021)	Y	Y	Y	Y	Y
Okkhornama (Talukder et al., 2021)	Y	N	N	Y	N
KU-BdSL (Jim et al., 2023)	Y	Y	Y	Y	Y
BdSL 49 (Hasib et al., 2023)	Y	N	N	Y	N
BdSL47 (Rayeed et al., 2023)	Y	N	Y	N	Y
BAUST Lipi (Hadiuzzaman et al., 2024)	Y	N	Y	Y	N
BDSLW-11 (M. M. Islam et al., 2022)	Y	N	N	Y	N
A15 (M. S. Islam et al., 2023)	Y	Y	Y	N	N
Sign-BD (Sams et al., 2023)	Y	Y	Y	Y	N
BdSLW60 (Rubaiyeat et al., 2025)	Y	Y	Y	Y	Y
BTVSL (Zeeon et al., 2024)	Y	Y	N	Y	N

6.2 Comparative Analysis of Bangla SLR Systems

BdSL recognition system outcomes are analyzed here, ranging from the input types and datasets to ML model implementation and performance. The common input modalities are cameras, Kinect sensors, data gloves, and accelerometers. Datasets contain images with single-hand, double-hand, or both-hand gestures. Among many common ML algorithms, such as pre-trained models like MobileNet-v1, CNN-LSTM, ResNet18, SVM, and KNN, CNN is found to be the most popular.

a. Highest Accuracy: In the case of alphabets and numeric, BdSL alphabets (99.86%) (Raihan et al., 2024) and BdSL numerals (99.8%) (Shamrat et al., 2021) achieve the highest accuracy, with the utilization of advanced CNN models and optimization methods. Combinedly, BdSL Alphabets and Numerals Classification also achieved near-perfect accuracy (99.99%) with ResNet18 (Podder et al., 2022).

b. Hybrid models also show high performance in BdSL recognition. Tasmere et al., 2020 proposed a deep CNN combined with GMM for preprocessing, which achieved the highest accuracy of 97.63% for BdSL digit recognition while M. M. Hasan and Ahsan, 2019 achieved 94.74% accuracy using HOG features and SVM.

c. However, these results are directly taken from the original papers- which use different datasets, feature engineering and classification models. Therefore, there is a chance of potential bias in the comparative performance analysis of these works. Future works may propose consistent dataset and framework to ensure fair comparison.

d. Dataset Size: The BdSL Dataset (23,786 images) and Ishara-Bochon/ Ishara-Lipi (1,075 and 1,005 images) are among the largest used. Basnin et al., 2021 used the largest dataset with 13,400 images, while Haque et al., 2019used the smallest dataset with 130 samples. The smaller datasets are found to be effective for BdSL recognition but lack diversity.

e. Preprocessing Techniques: Common preprocessing techniques include skin detection (YCbCr), background subtraction, grayscale conversion, histogram equalization, and GMM for moving object detection.

f. Methodology: CNN-based models are the most used ML methods for BdSL recognition, with different combinations like CNN-LSTM, Transfer Learning (VGG, InceptionV3), and ensembled network (CNN + Pose Estimation). Transfer learning with pre-trained models is applied to improve BdSL classification performance with small datasets. XAI techniques like SHAP and angular loss functions optimize ML interpretability and generalization.

g. Real-Time Applications: Multiple recent researchers, such as Khan et al., 2019 and Tasmere et al., 2020, prioritized developing a real-time BdSL detection system. These methods use edge devices (e.g., Jetson Nano), Raspberry Pi, or smartphone applications for this purpose. These devices ensure the regular accessibility of the BdSLR systems.

h. Generalization Challenges: Smaller datasets often yield lower ML performance (e.g., Haque et al., 2019), emphasizing the need for larger, more diverse datasets. Some research lacks comprehensive performance reports (Rony et al., 2018; Abedinc et al., 2023). The need for dataset diversity is addressed in a few studies, highlighting the challenge of inter-dataset evaluations (Youme et al., 2021).

Recent BdSL recognition systems mainly use similar data sources, and common research gaps for multi-hand inputs, real-time performance, and dynamic continuous sign deployment still exist. Despite high accuracy from CNN, SVM, and KNN models, challenges remain in computational efficiency and dataset quality. Further work is needed to optimize computational time for non-ideal datasets.

Regardless of progress, suitable models for hybrid systems in lightweight mobile applications remain scarce, and deep learning methods are limitedly explored. Current datasets have restricted vocabulary, often omitting digit and letter hand signs. Additionally, reliable real-time interpretation of a comprehensive Bangla sign lexicon is lacking.

6.3 Key Insights and Emerging Trends

The latest progress in ML techniques and the availability of large, annotated datasets have greatly improved BdSL recognition, enabling greater performance, efficiency, and real-time detection. Nonetheless, there are existing challenges regarding social acceptance of these systems, privacy, and language support.

6.3.1 Key Insights from existing literature

a. Advancements in Deep Learning and Real-Time Detection: Incorporating DL models like CNN and LSTM has drastically improved BdSL recognition and enabled real-time SLR for mobile devices. However, considering the dynamic real-world conditions, maintaining quality performance and recognition speed is a continual issue.

b. Improved Dataset Size and Diversity: Large annotated datasets (Hoque et al., 2021, Hasib et al., 2023) enhance precise BdSL recognition with diverse signs and relevant features. However, regional and community-specific differences pose challenges for generalizing BdSL systems.

c. Multimodal Integration: Multimodal data improves the recognition performance of BdSL systems, especially in challenging conditions like occlusion or poor lighting. However, real-time data processing is challenging for mobile devices with limited computational power and time.

d. Gesture Recognition Challenges: Continuous sign gesture recognition is still a challenging issue. Vision-based systems, though socially acceptable, require significant computational power, while sensor-based methods face social acceptability challenges due to their invasiveness, even though this is more effective.

e. Privacy and Security Concerns: For multimodal data-based or real-time BdSLR systems privacy concerns arise, which require strong encryption and secure data handling. Ethical issues around consent and data usage also require attention.

f. Social Acceptability and Cost Barriers: Social acceptability remains a barrier, as most users may find the BdSLR technology invasive and costly. Additionally, specialized devices often limit scalability and accessibility. These challenges work as barriers to developing affordable, scalable, and universally applicable BdSLR solutions.

6.3.2 Trend Analysis of BDSLR Systems

Technological advancements and the demand for practical solutions are driving the future of BdSL Recognition systems:

a. Shift Toward Deep Learning and Explainable AI: Transitioning to DL models is a significant step forward for BdSLR systems. Explainable AI holds a promising future for recognizing continuous SLR. As shown in Figure 2, a rising number of SLR systems are implementing DL and transformer learning. A similar trend is observed in the incremental implementation of explainable AI.

b. Multimodal Approaches for Enhanced Robustness: In BdSL systems, the combination of RGB camera, depth cameras, accelerometers, and other sensors are common. This multimodal approach improves gesture recognition by addressing issues like occlusion, noise, and poor lighting. However, real-time processing of multimodal data remains challenging for resource-limited devices. Figure 3 visualizes BdSL dataset types throughout the years. Here, the shift towards multimodal datasets is clearly visible.

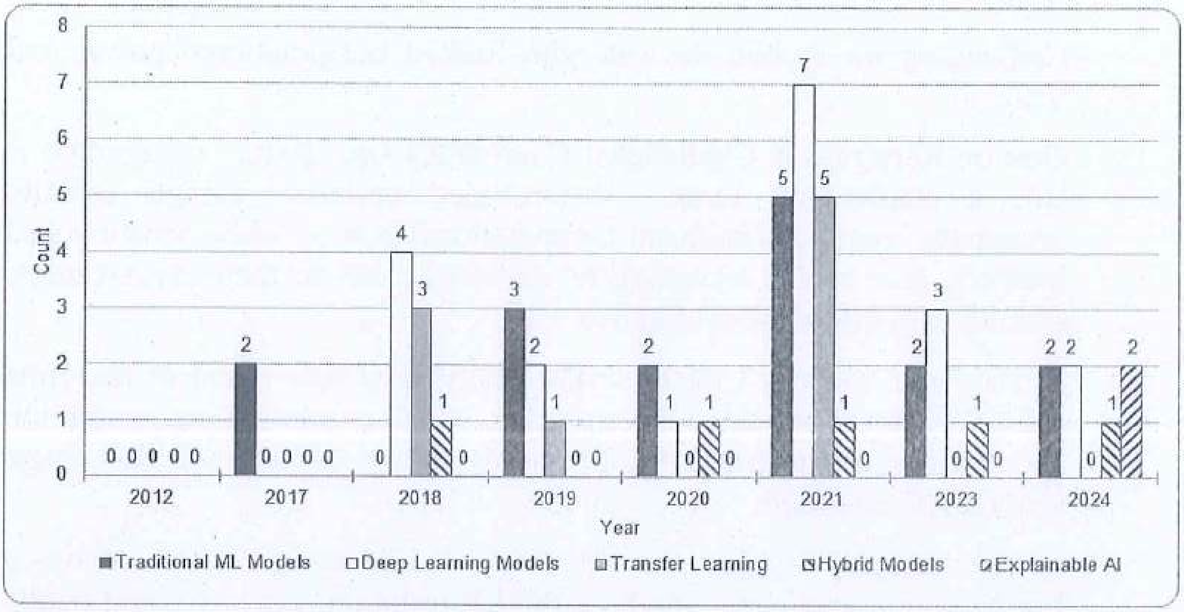

Figure 2: BdSLR model shifts toward Deep Learning and Transfer Learning

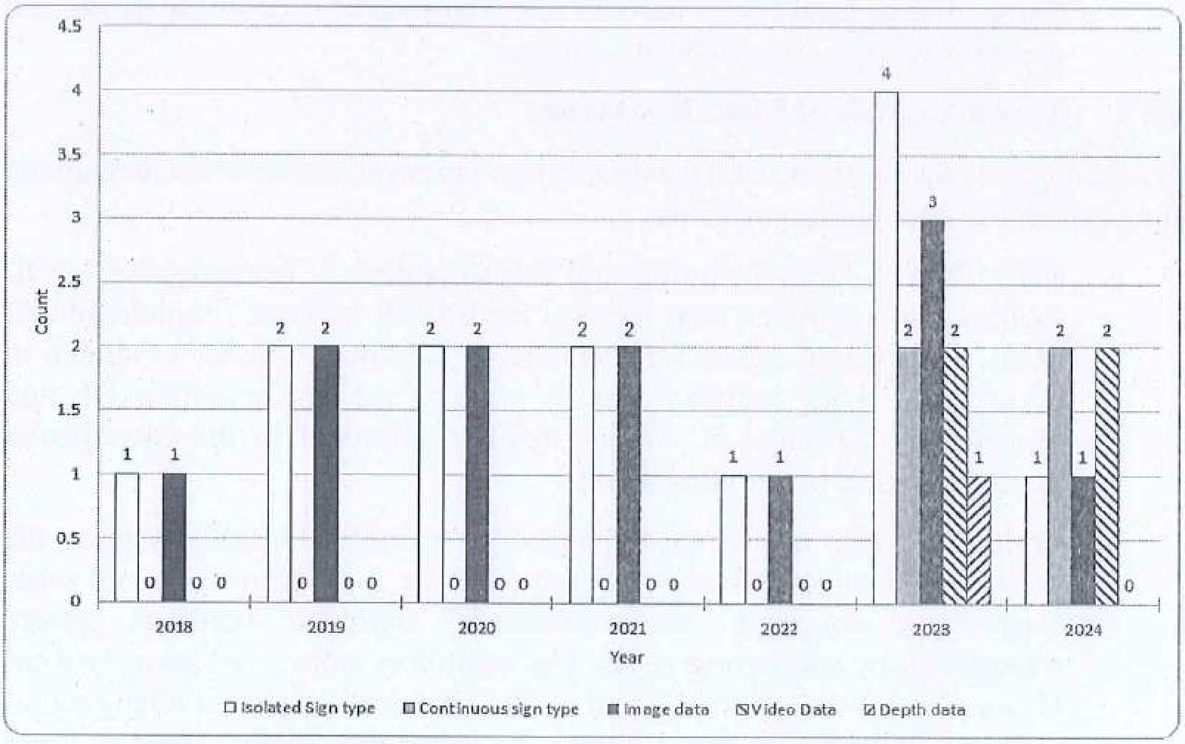

Figure 3: BdSL Dataset type shifts toward Multimodal data

c. Real-Time Performance: As mobile and IoT devices become more prevalent, optimizing BdSLR systems for real-time performance on these platforms is a key focus. Transformer models are being optimized to run efficiently on devices with limited processing power, like smartphones and IoT devices such as the Jetson Nano, making SLR more accessible in everyday environments. Figure 4 depicts the number of offline and real-time BdSLR systems with respect to years. Offline systems were common in the initial research; however researchers started to work with more real-time systems in recent years.

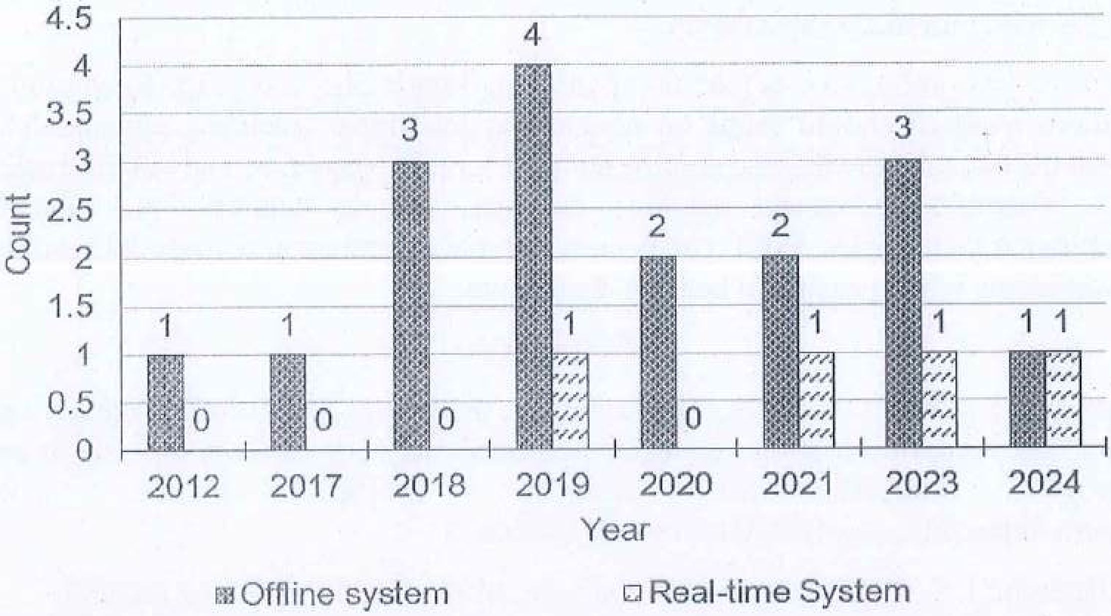

Figure 4: Number of Offline and Real-time BdSLR System

BdSL recognition systems have advanced with deep learning models, multimodal integration, and real-time capabilities. However, challenges in multilingual support, privacy, robustness, and social acceptability persist. Future researches should focus on inclusive datasets, improving model generalization across regional sign language variations, and addressing technical limitations in mobile and edge computing environments. Overcoming these challenges will enhance accuracy, accessibility, and social impact, empowering individuals with speech disabilities to communicate more effectively.

7 Conclusion and Future Directions

BdSL achieved notable progress due to the integration of machine learning methodologies. The technical progress of machine learning techniques presents promising avenues for breakthroughs in BdSL. Although CNN and RNN architectures constitute the foundation of current models, novel methodologies such as explainable AI, transformers, graph neural networks, GNNs, and reinforcement learning can markedly improve recognition accuracy and agility.

However, many challenges persist in obtaining more prominent accuracy, robustness, and scalability. This review emphasizes significant advancements in machine learning-based BdSL systems, datasets, and novel computational methodologies while also pinpointing essential areas for future exploration. A fundamental problem in BdSL research is the accessibility of high-quality, large-scale datasets. Mitigating data scarcity and class imbalance is essential for enhancing recognition accuracy and model generalization. The integration of self-supervised learning and few-shot learning methodologies may mitigate data scarcity, enabling the recognition models to generalize to novel and unidentified signs with minimum supervision.

This review outlines a roadmap for advancing Bangla Sign Language Recognition. Future research should focus on developing intelligent, scalable, and inclusive systems that adapt to diverse signing styles, user demographics, and environments. By integrating advanced machine learning, diverse datasets, and efficient deployment strategies, BdSL can become a highly accurate and accessible tool for individuals with speech and hearing disabilities.

References

Abedin, T., Prottoy, K. S. S., Moshruba, A., & Hakim, S. B. (2023). Bangla sign language recognition using a concatenated BdSL network. In *Computer Vision and Image Analysis for Industry 4.0* (pp. 76–86). CRC Press. https://doi.org/10.1201/9781003256106-7

Adeyanju, I. A., Bello, O. O., & Adegboye, M. A. (2021). Machine learning methods for sign language recognition: A critical review and analysis. *Intelligent Systems with Applications, 12*, 200056. https://doi.org/10.1016/j.iswa.2021.200056

Ahmed, S., Islam, M. R., Hassan, J., Ahmed, M. U., Ferdosi, B. J., Saha, S., & Shopon, M. (2019). Hand sign to Bangla speech: A deep learning in vision-based system for recognizing hand sign digits and generating Bangla speech. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3358187

Alam, I., Hameed, A., & Ziar, R. A. (2024). Exploring sign language detection on smartphones: A systematic review of machine and deep learning approaches. *Advances in Human-Computer Interaction, 2024*(1), 1487500. https://doi.org/10.1155/2024/1487500

Al-Qurishi, M., Khalid, T., & Souissi, R. (2021). Deep learning for sign language recognition: Current techniques, benchmarks, and open issues. *IEEE Access, 9*, 126917–126951. https://doi.org/10.1109/ACCESS.2021.3110912

Basnin, N., Nahar, L., & Hossain, M. S. (2021). An integrated CNN-LSTM model for Bangla lexical sign language recognition. In *Advances in Intelligent Systems and Computing,1309*, pp. 695–707). Springer. https://doi.org/10.1007/978-981-33-4673-4_57

Begum, S., & Hasanuzzaman, M. (2009). Computer vision-based Bangladeshi sign language recognition system. In *Proceedings of the 12th International Conference on Computer and Information Technology (ICCIT)* (pp. 414–419). https://doi.org/10.1109/ICCIT.2009.5407274

Das, S., Biswas, S. K., Chakraborty, M., & Purkayastha, B. (2021). A review on sign language recognition (SLR) system: ML and DL for SLR. In *2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT)* (pp. 177–182). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICISSGT52025.2021.00045

Das, S., Imtiaz, M. S., Neom, N. H., Siddique, N., & Wang, H. (2023). A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier. *Expert Systems with Applications, 213*, 118914. https://doi.org/10.1016/j.eswa.2022.118914

Elakkiya, R. (2021). Machine learning based sign language recognition: A review and its research frontier. *Journal of Ambient Intelligence and Humanized Computing, 12*(7), 7205–7224. https://doi.org/10.1007/s12652-020-02396-y

Hadiuzzaman, M., Ali, M. S., Sultana, T., Shafi, A. R., Miah, A. S. M., & Shin, J. (2024). *Baust lipi: A BdSL dataset with deep learning-based Bangla sign language recognition*. arXiv preprint. https://arxiv.org/abs/2408.10518

Haque, P., Das, B., & Kaspy, N. N. (2019). Two-handed Bangla sign language recognition using principal component analysis (PCA) and KNN algorithm. In *2019 2nd International Conference on Electrical, Computer and Communication Engineering (ECCE)* (pp. 1-4). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ECACE.2019.8679185

Hasan, M. M., & Ahsan, S. M. M. (2019). Bangla sign digits recognition using HOG feature-based multi-class support vector machine. In *2019 4th International Conference on Electrical Information and Communication Technology (EICT)* (pp. 1-5). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/EICT48899.2019.9068832

Hasan, S. N., Hasan, M. J., & Alam, K. S. (2021). Shongket: A comprehensive and multipurpose dataset for Bangla sign language detection. In *2021 International Conference on Electronics, Communications and Information Technology (ICECIT)*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICECIT54077.2021.9641137

Hasib, A., Eva, J. F., Khan, S. S., Khatun, M. N., Haque, A., Shahrin, N., Rahman, R., Murad, H., Islam, M. R., & Hussein, M. R. (2023). BDSL 49: A comprehensive dataset of Bangla sign language. *Data in Brief, 49*, 109329. https://doi.org/10.1016/j.dib.2023.109329

Hoque, O. B., Jubair, M. I., Akash, A.-F., & Islam, S. (2021). BdSL36: A dataset for Bangladeshi sign letters recognition. *Lecture Notes in Computer Science, 12628*, 71–86. https://doi.org/10.1007/978-3-030-69756-3_6

Hossen, M. A., Govindaiah, A., Sultana, S., & Bhuiyan, A. (2018). Bengali sign language recognition using deep convolutional neural network. In *2018 Joint 7th International Conference on Informatics, Electronics and Vision (ICIEV)* (pp. 369-373). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICIEV.2018.8640962

Islam, M. M., Uddin, M. R., Ferdous, M. J., Akter, S., & Akhtar, M. N. (2022). BdSLW-11: Dataset of Bangladeshi sign language words for recognizing 11 daily useful BdSL words. *Data in Brief, 45*, 108747. https://doi.org/10.1016/j.dib.2022.108747

Islam, M. S., Mousumi, S. S. S., Jessan, N. A., Rabby, A. S. A., Abujar, S., & Hossain, S. A. (2019). Ishara-Bochon: The first multipurpose open access dataset for Bangla sign language isolated digits. *Communications in Computer and Information Science, 1035*, 420–428. Springer. https://doi.org/10.1007/978-981-13-9181-1_37

Islam, M. S., Mousumi, S. S. S., Jessan, N. A., Rabby, A. S. A., & Hossain, S. A. (2018). Ishara-Lipi: The first complete multipurpose open access dataset of isolated characters for Bangla sign language. In *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)* (pp. 1-4). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICBSLP.2018.8554466

Islam, M. S., Joha, A., & Hasan, M. (2023). A15: Multi-view Bangla sign language dataset (MV-BSL) and continuous BSL recognition model.

Islam, R., Khan, M. R., Islam, M. Z., Alam, M. M., Hossain, M. S., Mehedi, M. H. K., & Rasel, A. A. (2023). Optimization of deep CNN-based Bangla sign language recognition using XGBoost classifier. In *2023 9th International Conference on Computer Technology Applications* (pp. 72–77).

Jim, A. A. J., Rafi, I., Akon, M. Z., Biswas, U., & Nahid, A. A. (2023). KU-BdSL: An open dataset for Bengali sign language recognition. *Data in Brief, 51*, 109797. https://doi.org/10.1016/j.dib.2023.109797

Jim, A. A. J., Rafi, I., Tiang, J. J., Biswas, U., & Nahid, A.-A. (2024). KUNet: An optimized AI-based Bengali sign language translator for deaf and dumb people. *IEEE Access, 12,* 155052-155063.

Khan, S. A., Joy, A. D., Asaduzzaman, S. M., & Hossain, M. (2019). An efficient sign language translator device using convolutional neural network and customized ROI segmentation. In *2019 2nd International Conference on Communication Engineering and Technology (ICCET)*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICCET.2019.8726895

Khatun, A., Shahriar, M. S., Hasan, M. H., Das, K., Ahmed, S., & Islam, M. S. (2021). A systematic review on the chronological development of Bangla sign language recognition systems. In *2021 Joint 10th International Conference on Informatics, Electronics and Vision (ICIEV)*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICIEVICIVPR52578.2021.9564157

Madhiarasan, M., & Roy, P. P. (2022). A comprehensive review of sign language recognition: Different types, modalities, and datasets. *arXiv preprint*. https://arxiv.org/abs/2204.03328

Mhatre, S., Joshi, S., & Kulkarni, H. B. (2022). Sign language detection using LSTM. In *2022 IEEE International Conference on Current Development in Engineering and Technology (CCET)*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CCET56606.2022.10080705

Moustafa, A., Rahim, M. M., Khattab, M. M., Zeki, A. M., Matter, S. S., Soliman, A. M., & Ahmed, A. M. (2024). Arabic sign language recognition systems: A systematic review. *Indian Journal of Computer Science and Engineering, 15*(1), 1–18.

Podder, K. K., Chowdhury, M. E. H., Tahir, A. M., Mahbub, Z. B., Khandakar, A., Hossain, M. S., & Kadir, M. A. (2022). Bangla sign language (BdSL) alphabets and numerals classification using a deep learning model. *Sensors, 22*(2), 574. https://doi.org/10.3390/s22020574

Rafi, A. M., Nawal, N., Bayev, N. S. N., Nima, L., Shahnaz, C., & Fattah, S. A. (2019). Image-based Bengali sign language alphabet recognition for deaf and dumb community. In *2019 IEEE Global Humanitarian Technology Conference (GHTC)*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/GHTC46095.2019.9033031

Rahman, M. A. (2012). Recognition of static hand gestures of alphabet in Bangla sign language. *IOSR Journal of Computer Engineering, 8*(1), 7–13. https://doi.org/10.9790/0661-0810713

Raihan, M. J., Labib, M. I., Jim, A. A. J., Tiang, J. J., Biswas, U., & Nahid, A.-A. (2024). Bengali-Sign: A machine learning-based Bengali sign language interpretation for deaf and non-verbal people. *Sensors, 24*(16), 5351. https://doi.org/10.3390/s24165351

Rastgoo, R., Kiani, K., & Escalera, S. (2021). Sign language recognition: A deep survey. *Expert Systems with Applications, 164*, 113794. https://doi.org/10.1016/j.eswa.2020.113794

Rayeed, S. M., Tuba, S. T., Mahmud, H., Mazumder, M. H. U., Mukta, S. H., & Hasan, K. (2023). BdSL47: A complete depth-based Bangla sign alphabet and digit dataset. *Data in Brief, 51*, 109799. https://doi.org/10.1016/j.dib.2023.109799

Renjith, S., & Manazhy, R. (2024). Sign language: A systematic review on classification and recognition. *Multimedia Tools and Applications, 83*(31), 77077–77127. https://doi.org/10.1007/s11042-024-18583-4

Rony, A. J., Saikat, K. H., Tanzeem, M., & Robi, F. M. R. H. (2018). An effective approach to communicate with the deaf and mute people by recognizing characters of one-hand Bangla sign language using convolutional neural-network. In *2018 4th International Conference on Electrical Engineering and Information and Communication Technology (iCEEiCT)*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CEEICT.2018.8628158

Rubaiyeat, H. A., Mahmud, H., Habib, A., & Hasan, M. K. (2025). BDSLW60: A word-level Bangla sign language dataset. *Multimedia Tools and Applications, 1–25.*

Sams, A., Akash, A. H., & Rahman, S. M. M. (2023). SignBD-Word: Video-based Bangla word-level sign language and pose translation. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICCCNT56998.2023.10306914

Sandjaja, I., Alsharoa, A., Wunsch, D., & Liu, J. (2024). Survey of Hidden Markov Models (HMMs) for Sign Language Recognition (SLR). In *2024 IEEE 7th International Conference on Industrial Cyber-Physical Systems (ICPS)* (pp. 1-6). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICPS59941.2024.10640040

Shamrat, F. M. J. M., Chakraborty, S., Billah, M. M., Kabir, M., Shadin, N. S., & Sanjana, S. (2021). Bangla numerical sign language recognition using convolutional neural networks. *Indonesian Journal of Electrical Engineering and Computer Science, 23*(1), 405–413. https://doi.org/10.11591/ijeecs.v23.i1.pp405-413

Shanta, S. S., Anwar, S. T., & Kabir, M. R. (2018). Bangla sign language detection using SIFT and CNN. In *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-6). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICCCNT.2018.8493915

Siddique, S., Islam, S., Neon, E. E., Sabbir, T., Naheen, I. T., & Khan, R. (2023). Deep learning-based Bangla sign language detection with an edge device. *Intelligent Systems with Applications, 18*, 200224. https://doi.org/10.1016/j.iswa.2023.200224

Subburaj, S., & Murugavalli, S. (2022). Survey on sign language recognition in context of vision-based and deep learning. *Measurement: Sensors, 23*, 100385. https://doi.org/10.1016/j.measen.2022.100385

Surjo, G. S., Ghosh, B. K., Alam, M. J., Mahamudullah, Razib, M., & Bilgaiyan, S. (2023). A comparative analysis between single & dual-handed Bangladeshi sign language detection using CNN based approach. In *2023 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-8). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICCCI56745.2023.10128583

Tabassum, A. T., Mahmud, I., Uddin, M. D. P., Emran, A. L. I., Afjal, M. I. B. N., & Nitu, A. M. (2020). Enhancement of single-handed Bengali sign language recognition based on HOG features. *Journal of Theoretical and Applied Information Technology, 15*(5).

Talukder, D., Jahara, F., Barua, S., & Haque, M. M. (2021). OkkhorNama: BdSL image dataset for real-time object detection algorithms. In *2021 IEEE Region 10 Symposium (TENSYMP)* (pp. 1-6). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/TENSYMP52854.2021.9550907

Tao, T., Zhao, Y., Liu, T., & Zhu, J. (2024). Sign language recognition: A comprehensive review of traditional and deep learning approaches, datasets, and challenges. *IEEE Access, 12*, 7503475060. https://doi.org/10.1109/ACCESS.2024.3398806

Tasmere, D., Ahmed, B., & Hasan, M. M. (2020). Bangla sign digits: A dataset for real-time hand gesture recognition. In *2020 11th International Conference on Electrical and Computer Engineering (ICECE)* (pp. 186-189). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICECE51571.2020.9393070

Uddin, J., Arko, F. N., Tabassum, N., Trisha, T. R., & Ahmed, F. (2017). Bangla sign language interpretation using bag of features and support vector machine. In *2017 3rd International Conference on Electrical Information and Communication Technology (EICT)* (pp. 1-4). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/EICT.2017.8275173

Youme, S. K., Chowdhury, T. A., Ahamed, H., Abid, M. S., Chowdhury, L., & Mohammed, N. (2021). Generalization of Bangla sign language recognition using angular loss functions. *IEEE Access, 9*, 165351–165365. https://doi.org/10.1109/ACCESS.2021.3134903

Zahid, H., Rashid, M., Hussain, S., Azim, F., Syed, S. A., & Saad, A. (2022). Recognition of Urdu sign language: A systematic review of the machine learning classification. *PeerJ Computer Science, 8,* e883. https://doi.org/10.7717/peerj-cs.883

Zeeon, I. E. M., Mohammad, M. M., & Adnan, M. A. (2024). BTVSL: A novel sentence-level annotated dataset for Bangla sign language translation. In *2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition (FG)* (pp. 1-10). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/FG59268.2024.10581873