A Detailed Analysis of Recent Advances in Automatic Sign Language Recognition

Sheetal Patil and Amol Kadam

September 12, 2024

A Detailed Analysis of recent advances in automatic sign language recognition

Sheetal Patil , Department of Computer Engineering Bharati Vidyapeeth (Deemed to be University)College of Engineering , Pune, India

Dr. Amol Kadam , Department of Bharati Vidyapeeth (Deemed to beUniversity)College of Engineering , Pune, India

ABSTARCT

This survey paper reviews the advancements in sign language recognition (SLR) and sign language translation (SLT) technologies. Both congenital and acquired Deaf and Hard of Hearing (DHH) people utilize sign language, a unique visual language that combines manual and nonmanual aspects for efficient communication. This paper explores various methods and models developed to enhance the accuracy and efficiency of SLR and SLT systems. Key techniques discussed include the use of deep learning frameworks such as Faster R-CNN, 3D-CNNs, and LSTMs, as well as hierarchical fusion models and skeleton-aware representations. Special attention is given to methods that address the challenges of precise action boundary detection, temporal cue learning, and robust key-point normalization. The paper also highlights the specific challenges encountered in different sign languages, such as the similarity of hand gestures in German sign language that differ only in lip shape. Through an analysis of these methods, the study seeks to offer a thorough grasp of the state-of-the-art in sign language technology.

Keyword: Sign Language, Machine learning, Deep learning, CNN,

[1]. INTRODUCTION

People who are Deaf and Hard of Hearing (DHH) by birth or who have acquired the language utilize sign language, a distinct visual language. It employs both manual and nonmanual components for visual communication. While arm motions, body posture, lip shape, eye contact and facial expressions are regarded as nonmanual aspects, the shape, orientation, location, and motion of the hands are considered manual elements. Sign language is not a literal translation of spoken language; rather, it has its own syntax, meaning structure, and linguistic logic. Repeated movements of the hands and body represent discrete meaning units. The World Federation of the Deaf estimates that there are 70 million DHH individuals worldwide and around 200 distinct sign languages. Enhancing sign language translation technologies can link the communication gap between DHH and non-DHH people. Previous work in sign language translation has primarily focused on sign language recognition, or the process of identifying sign language as similar glosses. SLT converts identified glosses into spoken language text rather than simply anticipating spoken language text from sign language movies. Glosses in sign language texts convey grammatical and semantic details related to tense, order, direction, and position, in contrast to writings in spoken language. Also, they might indicate whether a symbol is being repeated. Sign recognition can be classified into two categories: continuous and isolated. Segmenting isolated signs requires a lot of manual labor because each film represents a single gloss. Isolated sign recognition is the fine-grained recognition of individual sign motions. Full sign language movies are transformed into gloss sequences by continuous sign recognition, maintaining the original sign language's order. Three categories were established for the literature in this study: dataset, SLR type, and machine learning for detection.

[2]. LITREATURE REVIEW

To improve the acquisition of global visual semantic information, He et al. [1] employed the faster R-CNN model to detect and localize hand gestures in videos illustrating sign language. In order to meet the demanding accuracy standards for segmenting sign language videos, the researchers combined an encoder-decoder framework based on Long Short-Term Memory (LSTM) with a 3D Convolutional Neural Network (CNN) for sign language recognition (SLR). In order to capture visual information with varied granularities, Guo et al. [2] presented a hierarchical fusion model to explore precise action boundaries and learn temporal cues in sign language videos. At first, RGB features and skeletal descriptors were separately retrieved using a 3D-CNN architecture and a Kinect device. Subsequently, an adaptive clip summarization (ACS) framework was developed to automatically choose significant clips or frames of varying sizes. To learn features at the frame, clip, and viseme/signeme levels, multilayer LSTMs were applied. Finally, the target spoken text was generated using a query-adaptive model.

Gan et al. [3] suggested a skeleton-aware model that employed skeletons as a matched representation of human postures and separated the video into segments in order to fully use significant information from body postures and orientations. Kim et al. [4] proposed a key-point normalization strategy that utilized a neck-shoulder framework to standardize the placements of key points. This approach enhanced the resilience of their model. Subsequently, a transformer network received the normalized key points as its input. An important issue in the domain of German sign language is the occurrence of signs that share identical hand movements but have distinct lip shapes.

I. DATASET

TABLE : DATASET FOR SLT

Dataset Name	**Source/Citation**	**Description**	**Size/Volume**	**Equipment Used**	**Notes**
CopyCate Game	[6]	Data Collected from deaf children for scholastic adventure games using gesture recognition technology.	5,829 phrases over 4 phases, with 9 deployments	2 colored gloves (red and purple) on each hand	Each phrase has 4, or 5 signs from a dictionary of 22 signs (adjectives, objects, prepositions, subjects).
Multiple Dataset	[8]	ArSL datasets, 40 phrases and 80-word lexicon, repeated 10 times each.	N/A	Five sensor DG5-Vhand data glove with two Polhemus G4 motion trackers	Dataset 2 collected using a digital camera without gloves.
ArSL Dataset	[9]	Videos captured with digital cameras from deaf volunteers for training and testing models.	20 lexicons, 45 repetitions every word—20 for training, 18 for testing.	Digital cameras	Videos stored as AVI format, 25 frames per second, 640 × 480 resolution.
CORPUS-NGT	[7]	Nederlandes Gebarentaal NGT, accessible for researchers and studies.	72 hours of recordings	N/A	Often used words in daily correspondence.About one hundred local signers of various ages engaged in.
Oliveira et al.	[10]	Irish dataset took using hand shapes and hand movements.	468 videos	Two datasets for static and dynamic ISL recognition.	None
ISL-HS	[11]	Irish SL dataset with moving hand while signing each letter.	486 videos	N/A	6 persons performed ISL; videos with

					removed backgrounds provided. 23 labels, excluding j, x, and z letters.
Camgoz et al.	[12]	Turkish SL recorded with Microsoft Kinect v2 sensor, containing signs from various domains such as finance and health.	855 signs; 496 samples in health, 171 in finance, 181 everyday signs	Microsoft Kinect v2 sensor	Every sign was clicked by 10 persons and repetitive 6 times. Each user performed about 30-70 signs.
SMILE	[13]	Swiss German Sign Language (DSGS) assessment system for adult L2 learners, providing feedback on manual parameters like hand position, shape, and movement.	One hundred lexical words noted with nineteen adult L2 learners and eleven adult L1 signers.	N/A	N/A
Gebre et al.	[5]	British and Greek sign language datasets provided on Dicta-Sign corpus.	Recordings for 4 sign languages, 14 signers per language, 2 hours per language	N/A	British and Greek sign languages selected based on signer's skin color contrast with background. Achieved 95% F1 score accuracy.
Sahoo	[14]	Dataset of digital numbers (0-9) collected from 100 users.	5,000 images, each character repeated 5 times	Sony digital camera (16.1MP)	Image resolution resized to 200 × 300, images in JPEG format. Divided into training and testing groups.
RKS-PERSIANSIGN	[15]	Large dataset of Persian sign language (PSL) collected from 10 contributors.	10,000 videos (100 videos per PSL word)	N/A	Often used words in daily correspondence.
Joze and Koller	[16]	Dataset including about 1,000 signs.	1,000 signs, 200 signers, 25,000 videos	N/A	N/A

II. Wearable Sensor-Based Sign Language Recognition

Muhammad Al-Qurishi and Thariq Khalid's review (2014-2021) concluded that multimodal recognition (using both vision- and sensor-based channels) outperforms unimodal analysis. They highlight the importance of conceptual classification and offer a framework for researchers to address the advantages and disadvantages of different input modalities [52]. In [53], a quantitative overview of sign language recognition was provided, reviewing seventy-two studies (1991-2019) to identify common difficulties, best methods, and trends in wearable sensor-based systems. The review examined sensor configuration, research design, classification techniques, sign language variance, and performance measures, noting challenges and suggesting standardized data collection and evaluation processes. Zinah Raad Saee et al. reviewed sensory glove-based systems for sign language recognition, highlighting dataset size as a significant challenge for hand gesture identification. They analyzed literature from 2017-2022 to understand objectives, challenges, and recommendations in this field [54]. S kani proposed an automatic sign interpreter using gloves with wearable sensors to generate audio output from sign language, addressing the communication needs of the deaf [55].

III. Vision-Based Sign Language Recognition Systems

Boban Joksimoski and Eftim Zdravevski reviewed methods and challenges in sign language recognition (2010-2021), identifying key technological advancements in synthesis, visualization, and identification of sign language [56]. Farman Shah and Muhammad Saqlain Shah developed an automated system for Pakistani sign language using vision-based features and support vector machines (SVMs) with multiple kernel learning (MKL). They reported promising results compared to existing methods [57]. The authors of [58] reviewed hand gesture and sign language recognition methods, comparing various machine learning approaches for real-time systems. They highlighted the obstacles and evaluated the performance of different techniques to identify the most accurate and efficient methods. In [59], an overview of deep neural networks for continuous sign language recognition was provided. The proposed framework used bi-directional recurrent neural networks and deep convolutional neural networks, optimized for representation with limited data. They proposed a classification system for research articles and found that much of the work focused on static, isolated, single-handed signs. Their study aims to provide a roadmap for future research and facilitate knowledge development in this field [60] and a critical review of machine learning techniques for sign language recognition is provided, focusing on vision-based systems, feature extraction, and classification. It also offers a brief overview of sign language to speech translation, aiming to serve as an introduction to sign language interpretation and automatic hand gesture recognition. M. Madhiarasan's review [61] offers an extensive overview of sign language recognition, examining requirements, challenges, and advancements over the past decade. The paper identifies gaps in the field and provides recommendations for future research, discussing various sensing approaches and SLR architecture.

This review analyzes methods based on proposed classifications, highlights datasets from current projects, and suggests open research issues and directions [63]. Aamir Wali and Roha Shariq [62] review recent developments in sign language recognition, analyzing frameworks and algorithms. Their study classifies SLR into units such as words, sentences, or alphabets, and assesses datasets used in recent research. Ankita Wadhawan and Usha Mittal [64] proposed a dynamic sign language recognition system using Convolutional Neural Networks (CNNs). Their Indian sign language recognition model achieved 70% training accuracy on dynamic gestures, providing a basis for future research and improving model accuracy for better communication within the sign language community [65,67].

IV. DATA PRE-PROCESSING & FEATURE EXTRACTION

A. Sign Representation

Communication is aided by the use of grammatically structured manual and non-manual sign representations in sign language, which is a visual language. The form, orientation of the palm, movement of the fingers or hands, posture, tilting of the head, mouthing, and other aspects of the facial expression are examples of these representations. Eight example frames grouped in a temporal sequence were utilized by Tang et al. [17] to show the movement of two hands that were originally next to each other and subsequently separated. In [18], the signer's hand served as the representation for every motion in an experiment, and the shape of the hand sign was represented by a hand segmentation phase. Koller et al. [19] employed a double state to classify sixty hand shapes, whereas the rubbish class was assigned a single state. Using the left hand as the submissive hand and the right as the dominant hand, Zhou et al. [20] concentrated on right-handed signers. In their study of Bengali Sign Language, Hossen et al. [21] combined related sound alphabets into single signals to express 51 letters with 38 signs.

As explained in [22], a word in the Bahasa Indonesia language might have up to five signals associated with it. Independent Signed Indonesian (SIBI) representations for every word and prefix are consistently accomplished with a single sign. To represent 26 signs, Huang et al. [23] used 66 input units and 26 output units. Several research have compared hand and body features; the results in [24] show that, for sign language identification, body features outperform hand features by 2.27%. This could be due to the higher reliability and durability of body joints compared to hand joints.

B. Normalization and Filtering

Normalization is the process of normalizing input data according to predetermined principles in machine learning and deep learning to enhance the efficiency of AI technologies. Usually carried out during data pre-processing, this process can involve several statistical operations or media processing activities according on the machine learning architecture, sample variability, input format (text, image, or video), and the goal of the

automation tool. Modern Sign Language Recognition (SLR) techniques frequently incorporate normalization, and its benefits have been scientifically demonstrated [25]. There is a wide range of normalization strategies used in SLR investigations due to the many input modalities and aims. The majority of methods are visual and involve converting photos into common formats that algorithms can understand, frequently down to the pixel level in feature extraction and network training.
Image scaling and reshaping are basic SLR normalizing techniques, as shown by Kratimenos et al. [26] and others [27]. Garurel et al. [28] adjust frame sizes to match feature map dimensions by using mean values and standard deviations obtained during training. Cropping is another widely used technique that improves the quality of visual data by eliminating portions that are not part of the hands and face, which are necessary for sign language communication.. To accommodate for camera distance, cropped photos in [29] are normalized depending on average neck length. Based on the benchmark signer's major joint positions, [30] standardizes input from other signers. Using contour extraction, as in [31], pictures' backgrounds are eliminated while concentrating on hands. Frame downsampling lowers computational loads and standardizes clip quality for SLR methods that use video input.

C. Feature Extraction

A crucial stage in developing Sign Language Recognition (SLR) models is feature extraction, which has a big impact on how well the models train and how well they can distinguish between various signs and words. Features, which are obtained from unprocessed data, are frequently the locations of body parts—such as hands and faces—that are crucial for communicating in sign language. These characteristics are tallied using statistical procedures, and weights are allocated according on how discriminating they are. They allow neural models to learn the probability of associations with specific classes by being expressed as vectors in the latent space.

Tang et al. [17] found that considering the two hands as a single entity during feature extraction increased recognition accuracy. A analogous method in [18] solved difficulties with processing numerous image modalities by utilizing PCANet for feature extraction. By translating sensor input from both hands into feature vectors, Li et al. [32] demonstrated feature extraction without the necessity to reconstruct the exact shape, orientation, and placement of the hand.

via convolution layers to construct feature maps via image convolution, Camgoz et al. [33] employed 2D CNNs for spatial feature extraction. Different convolution and subsampling processes could extract spatial-temporal properties, based to observations from [34]. A Gaussian mixture model-hidden Markov model (GMM-HMM) was trained using manually extracted hand-crafted features from sign language films by Huang et al. [35].

3D CNNs were chosen for some study because of their ability to record temporal and spatial interactions. For example, the ResNet model produces representations of every video clip using a 3D CNN. In a similar line, [21] constructed a feature extraction neural network with multiple layers. While [37] uses a trained CNN as the feature extractor for an SVM, [36] applied a convolution layer to extract different input features.

From video sequences, Konstantinidis et al. [38] recovered a blend of skeletal and video features. Skeletal features encompassed the face, hand, and body, while video features included picture and optical flow. For the aim of extracting video features, the pre-trained VGG-16 network on ImageNet was utilized, while FlowNet2 was used for optical flow images.

D. FEATURE SELECTION

The most important phase in creating machine-learning models for sign language recognition (SLR) is feature selection. Through this procedure, the data is condensed into a more manageable set of pertinent attributes, which are then fed into machine learning algorithms [39]. Finding characteristics that greatly improve the algorithm's capacity to discriminate between various sign language classes is the primary goal in order to reduce computing demands and increase prediction accuracy. Various factors, including the method of choice, the volume and structure of the raw data, and the particular goals of the machine learning task, might influence the number of features that are chosen [40]. Researchers employ a variety of methodologies to assess and rank features based on their relevance, with the aim of selecting the most useful features for the model[41].

Feature selection techniques are generally divided into two main categories: **supervised** and **unsupervised**.

- **Filter Methods**: These techniques, including variance thresholding, correlation coefficients, and Chi-square tests, evaluate features based on intrinsic statistical properties to determine their relevance. For example, variance thresholds remove features with low variability, while correlation coefficients measure the relationship between features and the target variable.

- **Wrapper Methods**: These methods, such as forward feature selection and backward feature elimination, assess the performance of feature subsets by evaluating how well they work with a specific algorithm. Wrapper methods involve iterative processes to add or remove features based on their impact on model performance.
- **Embedded Methods**: Techniques such as LASSO regularization and random forest importance integrate feature selection directly into the model training process. LASSO regularization penalizes less important features, while random forests provide feature importance scores based on their contribution to the classification task.
- **Hybrid Approaches**: These methods combine elements of both Filter and Wrapper techniques to leverage their respective advantages. Hybrid approaches might use Filter methods for initial feature selection and Wrapper methods for final feature evaluation.

The choice of feature selection technique depends on the specifics of the project, including the type of classifier used, the characteristics of the data, and the goals of the machine learning task [41,67]. Researchers must carefully select the most effective methods to balance feature relevance with computational efficiency.

Category	Technique	Description	Applications	Key Findings
Machine Learning	Support Vector Machines (SVM)	Supervised learning model for classification tasks. Finds optimal hyperplane to separate data.	Sign language data classification, real-time performance comparison, large-vocabulary recognition, hybrid systems with HOG features.	Lower accuracy than DNN in some tasks, competitive with softmax for real-time performance, improved accuracy with hybrid systems.
Dimensionality Reduction	Principal Component Analysis (PCA)	Converts data into a new coordinate system with fewer dimensions for dimensionality reduction.	Fingerspelling recognition, reducing CNN feature map dimensions, handling high-variance Kinect data, enhancing MFCC features, feature extraction in SLR.	Effective for reducing dimensionality and overfitting, high accuracy with SAE+PCA, high classification rate with RPCA.
Sequential Data Modeling	Hidden Markov Models (HMM)	Statistical model for analyzing sequential data, revealing trends in motion patterns.	SLR since 1996, isolated hand gesture classification, dual and factorial HMMs, parameter optimization, parallel computing, language-based problem solving, input/output HMMs, hand position tracking, accuracy improvement, left/right HMM, combination with GMM, multi-camera analysis, HMM with PCA, HMM with RNN, calculation time reduction, combination with CRF, EM-based algorithm, continuous SLR model, GMM-HMM for trajectory and hand-shape recognition, temporal pattern recognition, combination with BLSTM-NN.	Requires extensive data for training, improved performance with dual and factorial HMMs, high accuracy for static signs with parameter-based improvements, successful hand position tracking with input/output HMMs, improved performance with hybrid systems, enhanced accuracy with RNNs.

TABLE DEEP LEARNING BASED REVIEW

Authors	Year	Key Focus	Techniques Used	Key Findings/Contributions
Jang et al.	2018	ASL word prediction	Deep Learning	Real-time ASL gesture recognition and word prediction system using deep learning.

Kumar et al.	2018	Gesture-based translation for speech-impaired	Computer Vision	system that translates hand gestures for people with communication impairments into text or speech.
Jha et al.	2018	Real-time sign language recognition	Convolutional Neural Networks (CNNs)	High accuracy in real-time sign language recognition for enhancing communication.
Singla et al.	2018	Review of ASL gesture recognition techniques	Review of Techniques	analysis of the benefits and drawbacks of cutting-edge hand gesture recognition methods.
Patel et al.	2019	Overview of sign language recognition and translation systems	Review of Techniques	Overview of different techniques for sign language recognition and translation, discussing their potential impacts.
Razzak et al.	2019	Real-time sign language recognition	LSTM and CNN	High accuracy system for real-time sign language recognition combining LSTM and CNN.
Kim et al.	2020	Word prediction for hard of hearing people	Deep Learning	Deep learning-based gesture recognition and word prediction system designed for hard of hearing individuals.
Singh et al.	2020	Review of hand gesture recognition systems	Review of Techniques	An analysis of the potential of different hand gesture detection systems for sign language communication.
Afzal et al.	2021	Sign language recognition real time	Multi-scale Hand Segmentation, CNNs	Multi-scale hand segmentation and CNN is used for achieving accuracy.

CONCLUSION

In this paper, we have examined various machine learning techniques applied to the domain of sign language recognition (SLR). These techniques play a crucial role in improving the accuracy, efficiency, and robustness of SLR systems. By utilizing machine learning models, we can effectively process and interpret the complex gestures involved in sign language, making communication more accessible. Our exploration has shown that different methodologies offer unique advantages, whether in terms of dimensionality reduction, feature extraction, or handling sequential data. The integration of these techniques has led to significant advancements in recognizing both static and dynamic signs, accommodating larger vocabularies, and facilitating real-time applications.The continuous development and refinement of these approaches have demonstrated their potential in enhancing SLR systems' performance. Hybrid models that combine the strengths of various machine learning techniques have shown particular promise in achieving higher accuracy rates and better generalization.Future research should focus on further integrating and optimizing these methodologies, exploring new hybrid approaches, and addressing any remaining challenges in real-time and large-scale sign language recognition. By doing so, we can continue to improve the efficacy and accessibility of SLR systems, ultimately fostering better communication and inclusivity for the deaf and hard-of-hearing community.

REFERENCES

1. S. He, "Research of a sign language translation system based on deep learning," in *Proceedings of the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM)*, Dublin, Ireland, 17–19 October 2019, pp. 392–396.

2. D. Guo, W. Zhou, A. Li, H. Li, and M. Wang, "Hierarchical recurrent deep fusion using adaptive clip summarization for sign language translation," *IEEE Trans. Image Process.*, vol. 29, pp. 1575–1590, 2019. [CrossRef]
3. S. Gan, Y. Yin, Z. Jiang, L. Xie, and S. Lu, "Skeleton-aware neural sign language translation," in *Proceedings of the 29th ACM International Conference on Multimedia*, Chengdu, China, 20–24 October 2021, pp. 4353–4361. *Electronics*, vol. 12, no. 2678, 2023.
4. S. Kim, C.J. Kim, H.M. Park, Y. Jeong, J.Y. Jang, and H. Jung, "Robust keypoint normalization method for Korean sign language translation using transformer," in *Proceedings of the 2020 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju Island, Republic of Korea, 21–23 October 2020, pp. 1303–1305.
5. B. G. Gebre, P. Wittenburg, and T. Heskes, "Automatic sign language identification," in *2013 IEEE International Conference on Image Processing*, Melbourne, VIC, 2013, pp. 2626–2630.
6. Z. Zafrulla, H. Brashear, P. Yin, P. Presti, T. Starner, and H. Hamilton, "American sign language phrase verification in an educational game for deaf children," *IEEE*, pp. 3846–3849, 2010, doi: 10.1109/ICPR.2010.937.
7. I. Z. Onno Crasborn and J. Ros, "Corpus-NGT. An open access digital corpus of movies with annotations of Sign Language of the Netherlands," Technical Report, Centre for Language Studies, Radboud University Nijmegen, 2008. [Online]. Available: http://www.corpusngt.nl.
8. M. Hassan, K. Assaleh, and T. Shanableh, "Multiple proposals for continuous Arabic sign language recognition," *Sensing Imaging*, vol. 20, no. 1, pp. 1–23, 2019.
9. A. Youssif, A. Aboutabl, and H. Ali, "Arabic sign language (ArSL) recognition system using HMM," *Int. J. Adv. Computer Sci. Appl.*, vol. 2, 2011, doi: 10.14569/IJACSA.2011.021108.
10. M. Oliveira et al., "A dataset for Irish sign language recognition," in *Proceedings of the Irish Machine Vision and Image Processing Conference (IMVIP)*, vol. 8, 2017.
11. I. Hernández, "Automatic Irish sign language recognition," M.S. thesis, Trinity College, University of Dublin, 2018.
12. N. C. Camgoz et al., "BosphorusSign: a Turkish sign language recognition corpus in health and finance domains," in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 2016, pp. 1383–1388.
13. S. Ebling et al., "SMILE Swiss German sign language dataset," in *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC)*, University of Surrey, 2018.
14. A. Sahoo, "Indian sign language recognition using neural networks and kNN classifiers," *J. Eng. Appl. Sci.*, vol. 9, pp. 1255–1259, 2014.
15. R. Rastgoo, K. Kiani, and S. Escalera, "Hand sign language recognition using multi-view hand skeleton," *Expert. Syst. Appl.*, vol. 150, p. 113336, 2020.
16. H. R. V. Joze and O. Koller, "MS-ASL: A large-scale dataset and benchmark for understanding American sign language," *arXiv preprint arXiv:1812.01053*, 2018.
17. A. Tang, K. Lu, Y. Wang, J. Huang, and H. Li, "A real-time hand posture recognition system using deep neural networks," ACM Trans. Intell. Syst. Technol., vol. 6, no. 2, pp. 1–23, May 2015.
18. S. Aly, B. Osman, W. Aly, and M. Saber, "Arabic sign language fingerspelling recognition from depth and intensity images," in Proc. 12th Int. Comput. Eng. Conf. (ICENCO), Dec. 2016, pp. 99–104.
19. O. Koller, H. Ney, and R. Bowden, "Deep hand: How to train a CNN on 1 million hand images when your data is continuous and weakly labelled," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 3793–3802.
20. Y. Zhou, G. Jiang, and Y. Lin, "A novel finger and hand pose estimation technique for real-time hand gesture recognition," Pattern Recognit., vol. 49, pp. 102–114, Jan. 2016.
21. M. A. Hossen, A. Govindaiah, S. Sultana, and A. Bhuiyan, "Bengali sign language recognition using deep convolutional neural network," in Proc. Joint 7th Int. Conf. Informat., Electron. Vis. (ICIEV) 2nd Int. Conf. Imag., Vis. Pattern Recognit. (icIVPR), Jun. 2018, pp. 369–373.
22. M. C. Ariesta, F. Wiryana, Suharjito, and A. Zahra, "Sentence level Indonesian sign language recognition using 3D convolutional neural network and bidirectional recurrent neural network," in Proc. Indonesian Assoc. Pattern Recognit. Int. Conf. (INAPR), Sep. 2018, pp. 16–22.
23. J. Huang, W. Zhou, H. Li, and W. Li, "Sign language recognition using real-sense," in Proc. IEEE China Summit Int. Conf. Signal Inf. Process. (ChinaSIP), Jul. 2015, pp. 166–170.
24. S. Wei, X. Chen, X. Yang, S. Cao, and X. Zhang, "A component-based vocabulary-extensible sign language gesture recognition framework," Sensors, vol. 16, no. 4, p. 556, Apr. 2016.
25. U. Farooq, M. S. M. Rahim, N. Sabir, A. Hussain, and A. Abid, "Advances in machine translation for sign language: Approaches, limitations, and challenges," Neural Comput. Appl., pp. 1–43, 2021.

26. A. Kratimenos, G. Pavlakos, and P. Maragos, "Independent sign language recognition with 3D body, hands, and face reconstruction," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Jun. 2021, pp. 4270–4274.
27. D. Bansal et al., "CopyCat: Using sign language recognition to help deaf children acquire language skills," in Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst., May 2021, pp. 1–10.
28. K. Gajurel, C. Zhong, and G. Wang, "A fine-grained visual attention approach for fingerspelling recognition in the wild," 2021, arXiv:2105.07625. [Online]. Available: https://arxiv.org/abs/2105.07625.
29. M. De Coster, M. Van Herreweghe, and J. Dambre, "Sign language recognition with transformer networks," in Proc. 12th Int. Conf. Lang. Resour. Eval. (ELRA), 2020, pp. 6018–6024.
30. L. Meng and R. Li, "An attention-enhanced multi-scale and dual sign language recognition network based on a graph convolution network," Sensors, vol. 21, no. 4, p. 1120, Feb. 2021.
31. P. P. Roy, P. Kumar, and B.-G. Kim, "An efficient sign language recognition (SLR) system using camshift tracker and hidden Markov model (HMM)," Social Netw. Comput. Sci., vol. 2, no. 2, pp. 1–15, Apr. 2021.
32. K. Li, Z. Zhou, and C.-H. Lee, "Sign transition modeling and a scalable solution to continuous sign language recognition for real-world applications," ACM Trans. Accessible Comput., vol. 8, no. 2, pp. 1–23, Jan. 2016.
33. N. C. Camgoz, S. Hadfield, O. Koller, and R. Bowden, "SubUNets: End-to-end hand shape and continuous sign language recognition," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 3075–3084.
34. G. Marin, F. Dominio, and P. Zanuttigh, "Hand gesture recognition with leap motion and kinect devices," in Proc. IEEE Int. Conf. Image Process. (ICIP), Oct. 2014, pp. 1565–1569.
35. J. Huang, W. Zhou, H. Li, and W. Li, "Sign language recognition using 3D convolutional neural networks," in Proc. IEEE Int. Conf. Multimedia Expo (ICME), Jun. 2015, pp. 1–6.
36. X. Jiang and Y.-D. Zhang, "Chinese sign language fingerspelling via six-layer convolutional neural network with leaky rectified linear units for therapy and rehabilitation," J. Med. Imag. Health Informat., vol. 9, no. 9, pp. 2031–2090, Dec. 2019.
37. H. B. D. Nguyen and H. N. Do, "Deep learning for American sign language fingerspelling recognition system," in Proc. 26th Int. Conf. Telecommun. (ICT), Apr. 2019, pp. 314–318.
38. D. Konstantinidis, K. Dimitropoulos, and P. Daras, "A deep learning approach for analyzing video and skeletal features in sign language recognition," in Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST), Oct. 2018, pp. 1–6.
39.] M. Kuhn and K. Johnson, Applied Predictive Modeling, vol. 26. New York, NY, USA: Springer, 2013.
40. B. Butcher and B. J. Smith, Feature Engineering and Selection: A PracticalApproach for Predictive Models, K. Johnson and M. Kuhn, Eds. Boca Raton, FL, USA: Chapman and Hall, 2020.
41. A. J. Ferreira and M. A. T. Figueiredo, ``Efcient feature selection lters for high-dimensional data," Pattern Recognit. Lett., vol. 33, no. 13, pp. 17941804, Oct. 2012
42. Y. Jang, H. Lee, H. Park, and S. Kim, "Word prediction system for American Sign Language based on gesture recognition," Sensors, vol. 18, no. 12, p. 4271, 2018.
43. R. Kumar, R. K. Jha, and S. Dutta, "Gesture-Based Sign Language Translation System for Speech Impaired and People," in 2018 4th International Conference on Computing Communication and Automation (ICCCA), pp. 191-196, IEEE, 2018.
44. R. K. Jha, R. Kumar, and S. Dutta, "Real-Time Sign Language Recognition Using Convolutional Neural Networks," in 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE), pp. 1-2, IEEE, 2018.
45. S. Singla, S. Bhalla, and S. Batra, "Hand Gesture Recognition for American Sign Language Translation: A Review," in 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-6, IEEE, 2018.
46. J. Patel, R. Singh, and S. Chakraborty, "Sign Language Recognition and Translation: A Review," in 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-5, IEEE, 2019.
47. M. A. Razzak, A. A. Shafie, and M. A. M. Noor, "Real-time Sign Language Recognition using LSTM and Convolutional Neural Network," in 2019 IEEE 15th International Colloquium on Signal Processing & Its Applications (CSPA), pp. 106-111, IEEE, 2019.
48. J. Kim, H. Lee, and S. Kim, "Gesture-based word prediction system for hard of hearing people," Electronics, vol. 9, no. 3, p. 438, 2020.

49. R. Singh, J. Patel, and S. Chakraborty, "A review of hand gesture recognition systems for sign language communication," in 2020 2nd International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), pp. 1-5, IEEE, 2020.
50. M. Afzal, A. A. Shafie, and M. A. M. Noor, "Real-Time Sign Language Recognition using Multi-scale Hand Segmentation and Convolutional Neural Network," in 2021 IEEE 17th International Colloquium on Signal Processing & Its Applications (CSPA), pp. 105-110, IEEE, 2021.
51. Y. Jang, H. Lee, and S. Kim, "Real-time Sign Language Translation using Attention-Based Convolutional Neural Network," Electronics, vol. 10, no. 5, p. 52, 2021.
52. MUHAMMAD AL-QURISHI, THARIQ KHALID Deep Learning for Sign Language Recognition: Current Techniques, Benchmarks, and Open Issues, IEEE, 2021:126917-126951
53. Karly Kudrinko, Emile Flavin, Wearable Sensor-Based Sign Language Recognition: A Comprehensive Review,IEEE Reviews, 2020,
54. ZINAH RAAD SAEED 1 , ZURINAHNI BINTI ZAINOL, A Systematic Review on Systems-Based Sensory Gloves for Sign Language Pattern Recognition: An Update From 2017 to 2022, IEEE,2023,123358-123377
55. S. M. H. Kani, S. Abdullah, U. S. Amanullah, and G. Divya, “Sign Language Recognition with Convolutional Neural Network”, IJRESM, vol. 6, no. 5, pp. 116–119, May 2023, Accessed: Jul. 26, 2024.
56. BOBAN JOKSIMOSKI 1, EFTIM ZDRAVEVSKI 1, PETRE LAMESKI1 Technological Solutions for Sign Language Recognition: A Scoping Review of Research Trends, Challenges, and Opportunities ,IEEE, 2022. 40979-40998
57. FARMAN SHAH1 , MUHAMMAD SAQLAIN SHAH, Sign Language Recognition Using Multiple Kernel Learning: A Case Study of Pakistan Sign Language, IEEE, 2021: 67548-6758
58. Janpreet Singh, Dalwinder Singh, A Comprehensive Review on Sign Language Recognition Using Machine Learning,IEEE,2022
59. Runpeng Cui, Hu Liu, A Deep Neural Framework for Continuous Sign Language Recognition by Iterative Training, IEEE 2018,1-12
60. Ankita Wadhawan1,· Parteek Kumar, Sign Language Recognition Systems: A Decade Systematic Literature Review,2021,786-813
61. M. MADHIARASAN1, A Comprehensive Review of Sign Language Recognition: Different Types, Modalities, and Datasets,2022,1-30
62. Aamir Wali1 · Roha Shariq, Recent progress in sign language recognition: a review,2023,1-20
63. ZINAH RAAD SAEED 1 , ZURINAHNI BINTI ZAINOL, A Systematic Review on Systems-Based Sensory Gloves for Sign Language Pattern Recognition: An Update From 2017 to 2022, IEEE,2023,123358-123377
64. Ankita Wadhawan,Usha Mittal, Indian Sign Language Recognition System for Dynamic Signs,2022
65. Deep Kothadiya 1,Chintan Bhatt, Deepsign: Sign Language Detection and Recognition Using Deep Learning,MDPI,2022,1-12
66. Kadam Amol, Shrivastava Anurag, Pawar Sonali K.,Patil Vinod H,Michaelson Jacob, Singh Ashish, “Calories Burned Prediction Using Machine Learning” Proceedings of International Conference on Contemporary Computing and Informatics, IC3I 2023, Pages 1712 - 1717
67. Kadam Amol, Vempaty, Lakshmi Namratha,Kumar Melam,Durga Prasnna, Patil Vinod H, “Tackling Climate Change with Artificial Intelligence” Proceedings of International Conference on Contemporary Computing and Informatics, IC3I 2023Pages 1704 - 1711