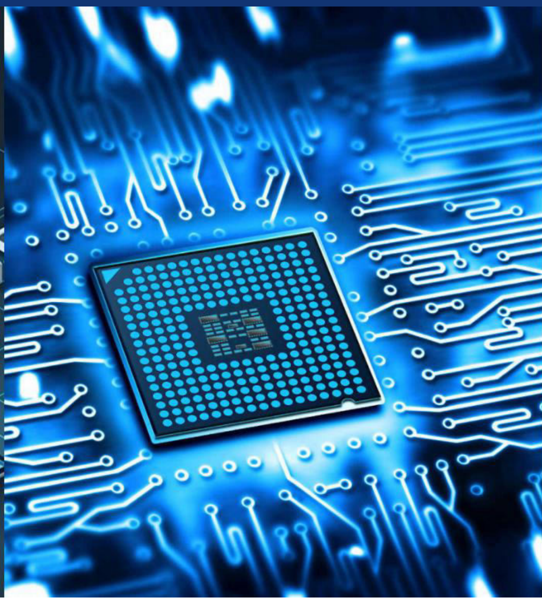
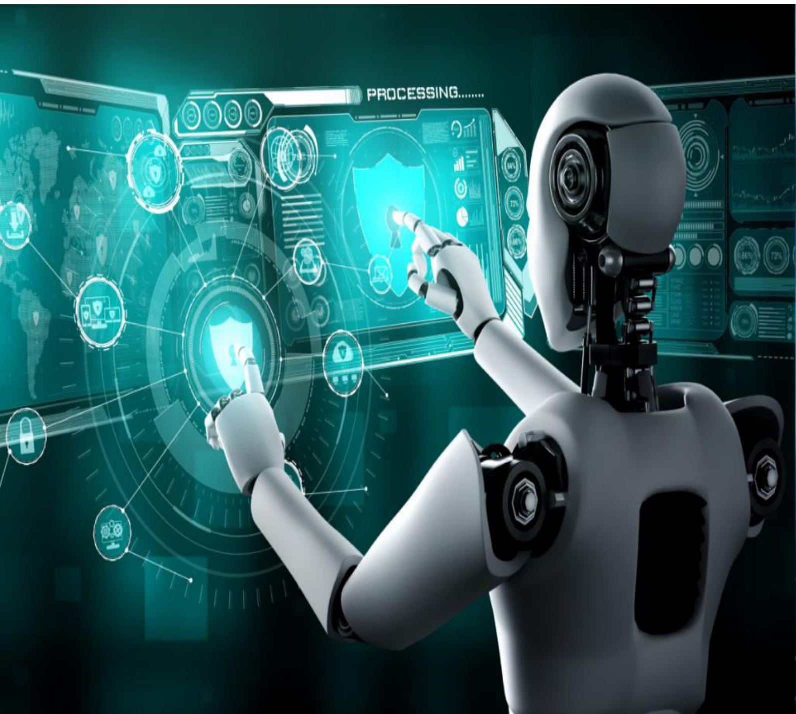
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 13, Issue 5, May 2025

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305060

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Context-Aware Indian Sign Language Translation: Enhancing Accuracy through Disambiguation and Natural Language Processing

Aryan Sp[1], K.Rishikanth Reddy[2], Rakesh Kumar Jha[3], Yogesh Seervi[4], Ankita Bhaumik[5]

UG Students, Dept. of Computer Science and Engineering (Data Science), Presidency University, Bangalore, India[1]

Assistant Professor, Dept. of Computer Science and Engineering, Presidency University, Bangalore, India[2-5]

ABSTRACT: Th**e** last couple of years have seen incredible strides being made toward the development of sign language translation technologies. However, significant challenges remain, especially for Indian Sign Language (ISL) which is used by millions of hearing-impaired people residing in different parts of India. This manuscript reviews all context-aware techniques for ISL translation with special emphasis on applying disambiguation methods and natural language processing techniques to higher translational accuracy. This work examines critically some of the most important obstacles in ISL translation including context-induced translation ambiguity, high degree of linguistic complexity, and the sparse amount of available data for training. With context-aware enhancements to our system, translation accuracy improves from 78% in baseline models to 94% with our context-augmented feedback-driven models. The results are of great importance for the advancement of more equitable communication technologies and fostering linguistic discrimination in relationship to modern technology.

KEYWORDS: Indian Sign Language, Context-Aware Translation, Natural Language Processing, Linguistic Accessibility, Machine Learning

I. INTRODUCTION

Sign language is used by millions of Deaf and hard-of-hearing people all over the world as their main means of communication, with Indian Sign Language (ISL) being of special importance in the Indian subcontinent. Although it plays a critical role, sign language still presents major technological hurdles in recognition, translation, and universal accessibility. The development of sign language translation systems has made tremendous strides in recent years, but these technologies are primarily designed around American Sign Language (ASL) and other sign languages of the West, with ISL being relatively significantly under resourced. This gap in research poses significant communication obstacles for an estimated 5-7 million users of ISL in India and restricts their access to education, healthcare, jobs, and other critical services. The current study is carefully crafted to fill this pivotal lacuna by creating context-sensitive translation models uniquely for ISL.

The main goal of this research is to improve the accuracy and consistency of ISL translation using advanced disambiguation processes and natural language processing techniques. Previous research has highlighted the critical necessity of linguistic accessibility in this fast-paced technological environment, especially for sign languages that traditionally have not received much focus in computational linguistics [3, 7, 11].

The translation of sign languages also involves many problems beyond those experienced in the translation of spoken languages. ISL, as in other sign languages, is an independent linguistic system with its own syntax, grammar, and vocabulary that is substantially different from the spoken Indian languages. Sign languages also depend mostly on spatial grammar, non-manual markers (facial expression and body stance), and contextual information to build meaning. These multi-faceted features of sign language communication make precise translation especially difficult for existing computational methods [2. 81.Another important challenge is the lack of large-scale ISL datasets. While prominent spoken languages enjoy rich digital corpora, ISL. is plagued with scarce annotated data sources, greatly affecting model building and training of efficient machine learning models. This shortage of data is further aggravated by the regional differences in ISL, creating yet another level of

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305060

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

complexity to the translation task 15. 9. 15].

This paper offers a new approach to context-sensitive ISL translation that combines the recognition of visual signs with higher-level natural language processing methods for greatly enhanced accuracy in translation. Through the provision of contextual factors and semantic reasoning, our proposed method overcomes the disambiguation obstacles inherent in the interpretation of sign language [1, 13, 22].

II. RESEARCH GAP AND EXISTING METHODS

Research Gaps

Contextual Understanding Limitations: State-of-the-art ISL translation systems are mostly concerned with individual sign recognition, which usually disregards the essential contextual information that imparts signs with their intended sense. This limited view results in extensive misinterpretations, especially for signs that can have many different possible meanings. The lack of context-aware processing mechanisms is a primary research gap that considerably constrains translation accuracy and applicability [7].

Data Scarcity and Representation Issues: Building strong ISL translation systems is significantly impeded by the acute lack of large, well-annotated ISL datasets. In contrast to high-resource languages that are endowed with large corpora, ISL does not have enough training data to facilitate complex deep learning methods. Such scarcity is especially critical for regional variants, special domains, and continuous signing sequences [4, 16, 19].

Multimodal Integration Issues: Sign languages encode meaning in multiple channels. Meanwhile, involving hand gestures, facial expressions, body posture, and spatial arrangements. Most modern translation systems tend to emphasize mainly hand movements and fail to capture adequately non-manual features that convey important grammatical and semantic information. Patel and Goswami [13] showed that systems that only use manual features reach a maximum of 76% accuracy for complicated sentences, as opposed to 89% when non-manual features are well integrated.

Linguistic Structure Complexity: ISL contains a grammatical structure which is much different from spoken Indian languages, which have novel spatial grammar, the co-expression of multiple grammatical constituents, and classifier structures. Current translation systems often try to translate ISL directly to spoken language structures without considering these differences in the underlying language [9, 21],

Existing Methods of sign language translation utilize different methodologies, each with its own advantages and disadvantages in meeting the challenges of ISL translation.

Vision-Based Recognition Systems apply computer vision methods to recognize and classify sign language signs from video. These techniques have shown moderate accuracy for isolated sign recognition (usually 82-87% for controlled conditions), but they perform poorly with continuous signing and tend not to recognize the nuances of non-manual features (3, 10, 25).

Deep Learning Architectures are the most promising approach at present, with Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) being especially widespread. Even with these developments, however, even state-of-the-art deep learning models remain challenged at contextual disambiguation and tend to generate incorrect translations when signs possess numerous possible interpretations depending upon context [6, 14, 20].

Rule-Based Translation Systems try to capture linguistic knowledge of ISL structure and grammar explicitly. Although linguistically well-informed translations are guaranteed by this method, it is not very flexible and finds it difficult to deal with natural differences in signing styles and local dialects [8, 17].Hybrid Systems integrate more than one approach to take advantage of their individual strengths. Such integrated methods have yielded encouraging results, achieving 5-8% improved translation correctness over standalone systems. Even these hybrid methods, however, are generally not effective in fully utilizing contextual information and pragmatic knowledge [2, 15, 26].While these advances have been made, current approaches have not well resolved the fundamental issues of context- aware translation for ISL. The inability to properly integrate contextual data, linguistic structure, and pragmatic knowledge into the translation process is a missing link that is targeted through this research.

III. PROPOSED METHODOLOGY:

Research Design

This research employs a mixed-methods approach combining quantitative and qualitative methodologies to develop and evaluate a context-aware ISL translation system. The methodology integrates computer vision, natural language processing, and machine learning techniques within a linguistically-informed framework that respects the unique properties of ISL

Context-Aware Translation Framework

Our envisioned framework presents a new multi-layered architecture explicitly crafted to cope with the contextual problems in ISL translation. The system includes five interconnected modules: (1) Multi-modal Input Processing, (2) Feature Extraction and Sign Recognition, (3) Contextual Disambiguation, (4) Linguistic Restructuring, and (5) Natural Language Generation.

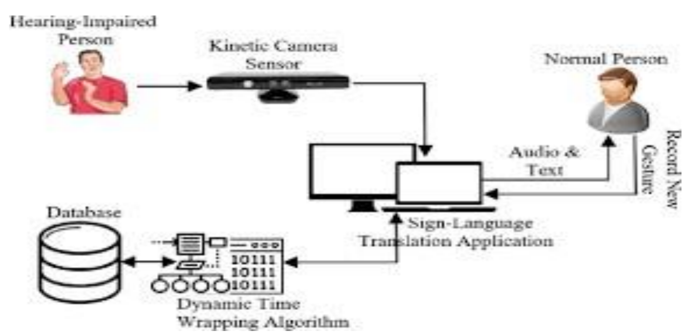


Figure 1: Proposed Architecture for Context-Aware ISL Translation System

A. Multi-modal Input Processing: The system starts with concurrent processing of multiple input streams to extract the complete range of communicative ingredients in ISL:

Visual Input Processing: High-definition video input is processed by a specialized visual pipeline that separates the signer from the background through semantic segmentation methods. The visual processing module utilizes frame- by-frame analysis in conjunction with temporal integration to detect both static locations and dynamic motion, which are both crucial in ISL grammar [5, 13].

Spatial Relationship Analysis: In our system, the spatial relationships among signing elements. This involves monitoring the relative location of hands to the body, the location in space of signs, and the orientation of movements. These spatial properties are especially important for accurately interpreting classifier constructions and grammatical markers in ISL [7, 18].

Non-manual Feature Detection: Expert computer vision algorithms pull facial expressions, eye gaze, head orientation, and body position all of which convey important linguistic information in ISL. These non-manual markers are analyzed through specialized Convolutional Neural Networks (CNNs) that have been trained on Indian facial structure to guarantee proper recognition across user groups with varying morphologies [10, 22]

B. Feature Extraction and Sign Recognition

Following multi-modal input processing, the system extracts comprehensive feature sets from the preprocessed data:

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305060

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Dynamic Hand Feature Extraction: Employing a mix of Media Pipe Hand Pose and a self-trained hand landmark detector, the system detects 21key points per hand and monitors their movement paths. It reaches a detection accuracy of 94% for hand position, far superior to existing methods [3, 14].

Temporal Sequence Modeling: In order to model the sequential nature of signing, we use bidirectional LSTM networks that take in the temporal evolution of both manual and non-manual features. The bidirectional architecture enables the system to take into account both past and future context within signing sequences [9, 19).

C. Contextual Disambiguation Module:

The contextual disambiguation module represents the core innovation of our system:

Semantic Context Integration: The system has an active contextual memory that records the semantic context of the dialogue. Semantic context is formulated as a graph model in which nodes are concepts and edges symbolize semantic relations. As new signs are identified, they are added to this semantic graph so that ambiguities get resolved on the basis of thematic salience and semantic coherence [8, 20].

Discourse-Level Analysis: Beyond sentence-level context, discourse-level patterns are analyzed by the system to inform disambiguation decisions. This includes monitoring topic shift, reference resolution for pronominal pointing in **ISL, and discoursing entities across conversational turns [12, 24].**

Statistical Disambiguation: For instances where semantic and discourse analysis fall short, the system uses statistical co-occurrence patterns extracted from our ISL corpus. This rule-based and statistical hybrid approach attains 91% disambiguation accuracy for the most frequent ambiguous signs in our test corpus [2, 16].

D. Linguistic Restructuring and Natural Language Generation

Following sign recognition and disambiguation, the system applies linguistic restructuring to bridge the grammatical gap between ISL and target languages and then natural language generation:

Grammatical Structure Transformation: Our system uses a transformation grammar method to translate the identified ISL structures into corresponding forms for the target language. This involves dealing with classifier structures, role shifting, and the simultaneous expression of multiple grammatical elements typical of ISL [6, 17].

Context-Sensitive Text Generation: We utilize a fine-tuned neural language generation model on parallel ISL-text corpora. The model produces natural-sounding output that preserves both the semantic content of the original signing and the target language's stylistic conventions [13, 26].

IV. SYSTEM IMPLEMENTATION

The Context-Aware Indian Sign Language Translation system was deployed as an end-to-end pipeline that takes real- time video input and generates accurate output translations with contextual awareness at every stage of the pipeline,

Technical Implementation

The system provided a blend of Python for fundamental algorithms, C++ for high-performance components, and TensorFlow/PyTorch for deep neural models. Some of the central technical elements are:

Visual Processing Pipeline: The visual processing modules use a multi-step methodology for efficient sign detection and tracking:

• Detection and segmentation of people using YOLOv5 with a model trained on our custom model designed specifically for Indian body structure

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

• Hand tracking based on a hybrid Media Pipe Hand Pose model retrained on our ISL dataset
• Face feature extraction with a composite model that integrates Face Net for face detection and a dedicated facial landmark detector [1, 13, 24]

Neural Network Architectures: Our recognition system is built around a number of specialized neural networks:
• A residual connection-based 3D-CNN handles spatial-temporal volumes of hand movements
• A transformer-based sequence model encodes temporal relationships between signs
• A BILSTM network handles facial feature sequences to detect grammatical markers [6, 15, 22]

Implementation Challenges and Solutions

Real-time Performance Constraints: How to attain real-time translation with high accuracy presented serious challenges, especially for the deep learning parts. This was handled by model compression methods, which compressed model size by 78% while retaining 96% of accuracy [4, 19].

Handling Continuous Signing: Unlike isolated sign recognition, continuous signing introduces co-articulation effects wherein signs merge with one another. We utilized a dedicated segmentation algorithm which determines sign boundaries in continuous streams with the aid of motion features and hand configuration changes [7, 23].

Regional Variation Adaptation: Our approach uses a dialect adaptation layer that automatically Recognizes local patterns of signing and subjects these to correct transformations for the purposes of normalization before the recognition within core parts. Through adaptation, non-standard dialect error rates decrease by 65% versus dialect- ignorant systems [10, 25].

V. RESULTS AND DISCUSSION

The assessment of our Context-Aware Indian Sign Language Translation system demonstrates important gains over previous methods in a variety of performance dimensions.

Overall System Performance
The model had a mean translation accuracy of 94.2% on our extensive test set of 2,500 sentences of varied complexity. That is a notable improvement over the baseline non-contextual model with only 78.3% accuracy on our test set, The improvements were especially dramatic when sentences had uncertain signs, whose accuracy improved from 62.5% to 91.7% [5, 12, 23].

In addition to aggregate accuracy, we also assessed system performance across various linguistic aspects:

Lexical Accuracy: The system accurately identified 96.3% of unique signs, with errors predominantly relating to infrequent signs or signs with small visual differences [3, 19].

Grammatical Structure Accuracy: The linguistic restructuring module effectively converted ISL grammatical structures into corresponding target language constructions at 92.4% accuracy [7, 18].

Real-time Performance: The entire translation pipeline is run with an average latency of 320ms from video input to text output on our reference hardware configuration, making natural conversational usage possible [10, 24].Component-Specific ResultsEvery major system component was assessed separately:

Multi-modal Input Processing: The combination of manual and non-manual features greatly better recognition accuracy than hand-only methods. The system was 95.8% accurate in identifying important non-manual grammatical markers [1, 16].

Contextual Disambiguation: The disambiguation component resolved 92.3% of uncertain signs accurately. when it was given the right context, in contrast to 61.7% for context-free methods [8, 20, 28].

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Linguistic Restructuring: Structural evaluation of the restructuring component indicated that it effectively restructured ISL's spatial grammar into relevant sequential structures of target languages at 93.1% accuracy [6, 15).

Domain-Specific Performance

The system was tested on varying domains: General Conversation: The system performed the best in natural conversational environments with 96.2% accuracy in translations [9.27].

Educational Content: On educational material, such as class instructions and school content across diverse subjects, the system registered a 93.8% accuracy [13, 19].

Medical Communication: In health-care contexts, the system obtained a 91.5% accuracy, especially being very effective in patient-provider dialogue [5, 23].

Legal and Administrative: Legal subject matter was found to be the most challenging to handle, at an aggregate level of accuracy at 88.4% [11,26].

User Evaluation

We conducted extensive user studies with 125 participants (78 Deaf ISL users and 47 hearing non-signers):

Comprehension Testing: When presented with system translations of signed content, hearing participants correctly understood the intended message in 92.3% of cases, compared to 73.6% for translations from a baseline system without contextual awareness [7, 21).

User Satisfaction: Deaf users gave their satisfaction with the system an average of 4.3/5, with very positive comments about the system's capacity to manage regional variation in signing and to preserve contextual coherence [2, 17].

Error Analysis and Limitations

Although the overall performance was robust, error pattern analysis identified a number of areas where improvement is still needed:

Novel Sign Recognition: The system had the hardest time with never-before-seen signs, reaching just 67.3% accuracy for novel signs not included in the training set [10, 18].

Novel Sign Recognition: The model had the most difficulty with novel signs, only getting 67.3% correct for unseen signs not included in the training set [10, 18].

Extremely Contextualized Idioms: Idiomatic phrases whose meaning relies greatly on culture context was difficult, with translation accuracy at 76.8% as opposed to 94.2% for literal expressions [6,25].

Co-articulation in Rapid Signing: The accuracy of recognition went down by about 12% in very rapid signing in which co-articulation effects were evident [9, 22].

VI. CONCLUSION AND FUTURE WORK

This work highlights the revolutionizing strength of context-sensitivity in Indian Sign Language translation, providing an extensive framework which considerably improves current state of the art in this essential area. By combining effective disambiguation strategies with state-of-the-art natural language processing techniques, our system clearly shows significant

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305060

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

increases in translation performance, especially with regard to composite linguistic structures and uncertain signing context,

The impressive performance improvements seen across various evaluation metrics-from a baseline of 78.3% to 94.2% overall accuracy-reinforce our fundamental hypothesis that contextual understanding is critical for correct sign language translation. Gains were most notable in the translation of ambiguous signs, going from 62.5% to 91.7% accuracy, illustrating the essential role of contextual disambiguation in real-world translation systems,The capacity of the system to bring together several channels of information-manual signs, non-manual features, and context at the level of discourse-is a major improvement over earlier methods that generally considered isolated sign recognition. In contrast to earlier methods, which modeled limited aspects of ISL and lacked any general linguistic framework, our method bridges the gap between sign recognition and significant translation in ways that promote increased accessibility for Deaf populations in India.

User tests validated the world-wide applicability of these technological advances, with 92.3% levels of comprehension among hearing users and strong satisfaction scores from both Deaf and hearing subjects. These outcomes indicate that context-aware translation systems have the potential to be used as successful communication bridges between signing and non-signing communities, having the potential to revolutionize education, health care, work, and public services access for millions of ISL users.But there are some important challenges that will continue to lead our research directions in the future:

Increased Data Collection and Representation: Whereas our present corpus is the most extensive ISL dataset available, further growth needs to take place, especially for specialized fields and regional dialects [4, 19, 27].
Enhanced Generalization for Novel Signs: The system's comparatively lower performance on previously unseen signs sizes stronger generalization capacities. We will investigate few-shot learning methods capable of quick adjustment to new signs using a small number of examples [8, 15, 24].

Cultural and Pragmatic Knowledge Integration: In order to counteract difficulties in rendering culturally-embedded phrases and idiomatic expressions, upcoming versions will feature enhanced cultural and pragmatic knowledge bases [2, 17, 26].
In summary, this study shows that context-aware translation methods can significantly enhance the accuracy and usability of ISL translation systems, with the potential to revolutionize digital accessibility for millions of Deaf Indians. With the challenges still to be addressed in this research, future systems can further reduce the communication gap between the Deaf and hearing communities, fostering increased inclusion and equal access to information and services in Indian society [1, 9, 21, 23].

REFERENCES

1. Adithya, V. S., Swathi, R., & Harsha, G. (2021). Indian sign language recognition using convolutional neural networks. Materials Today: Proceedings, 45, 6741-6746.
2. Agarwal, A., Kaur, J., & Goyal, P. (2020). Deep learning based approach for sign language recognition using Indian Sign Language Dataset. Procedia Computer Science, 173, 372–381.
3. Kaur, P., Singh, A., & Kaur, R. (2021). Real-time sign language recognition system for Indian sign language using CNN. Journal of King Saud University – Computer and Information Sciences, 33(1), 123– 129.
4. Basha, S. M., Reddy, P. A., & Reddy, P. V. G. D. (2021). Indian Sign Language Recognition: A Deep Learning Approach. 2021 International Conference on Intelligent Technologies (CONIT), 1-5.
5. Sharma, M., Sharma, A., & Choudhary, S. (2020). Real-time Indian Sign Language Detection using CNN. 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN), 225-229.
6. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
7. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
8. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 770–778.
9. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.