Journal of Information Systems Engineering and Management
2025, 10(22s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Secure Hand Landmark-Enhanced CNN: Advancing Safe Sign Language Interpretation

Manpreet Singh Bajwa[1], Rakesh Kumar Saxena[2]

[1]Research Scholar, FCE, Poornima University, Jaipur, Rajasthan , India manisinghbajwa@gmail.com
[2]Professor, FCE, Poornima University, Jaipur, Rajasthan , India saxenark06@gmail.com

ARTICLE INFO

Received: 24 Dec 2024

Revised: 06 Feb 2025

Accepted: 20 Feb 2025

ABSTRACT

Introduction: To bridge the gap between individuals with problems concerning listening and hearing, the sign language becomes an effective tool. Sign language is a complex structured language with variety in gesture, expressions, hand signals and movements to completely transfer the content correctly which becomes a challenge for automatic detection and interpretation accuracy is important. Advancements in Machine learning particularly in area of deep learning after Convolutional Neural Networks (CNNs), has open the avenues to improve accuracy and keep in mind the computer vision area for better object detection and made the accuracy a point to be keep forward for Sign language.

Objectives: This article proposes a novel approach to address the problem of interpreting American Sign language in alignment with security aspect of this complete CNN based model.

Methods This approach is a blended version of hand landmark characteristics and CNN feature extraction to improve the precision, accuracy and security of recognition of Indian sign language. The interesting point in this proposed method is that the finer details often missed by CNN will not be missed with this proposed CNN based technique for sign language. At the time of review and analysis, the proposed CNN has outperformed all other model for clarity in detecting the sign language hand gestures.

Results: Keeping in mind the facial expression, lightning, signer orientation and other background factors, the proposed CNN is resilient against variations in background

Conclusions:By focusing on accuracy and security of sign language interpretation our proposed model is more reliable and dependent as a communicating medium for the deaf and hard-of-hearing community.
Keywords: Hand Landmarks, CNN, Sign Language Interpretation, ASL, Deep Learning.

INTRODUCTION

This article is proposing an innovative CNN architecture to be reliable and accurate with secure platform for sign language interpretation. By integrating hand landmark data with advanced CNN capabilities, our approach aims to enhance both the precision and effectiveness of ASL recognition, addressing potential vulnerabilities and ensuring the trustworthiness of these systems in various real-world applications. The focus on promoting inclusion and reducing communication gaps leads to innovation in sign language interpretation, especially automatic systems. There are situations when the deaf and hard of hearing people face problems due to limited accessibility to the human sign language interpreter which then results in inability to indulge in social involvement, daily interactions, and access to essential services like healthcare, education, and legal counsel. The primary objective is to utilize technology to support human efforts so that the people who use sign language can communicate efficiently and independently irrespective of their location and time. To address these issues a unique opportunity is offered by technological advancements in computer vision and deep learning. Automatic sign language interpretation systems aim to provide a fast, affordable, and reliable means of communication, overcoming the limitations set by the lack of skilled interpreters and the logistical challenges of scheduling interpretation services. However, as the system are deployed, it is important to make sure its security and reliability to protect users' privacy and trust. In addition to enhancing accessibility, the aim is to empower the deaf and hard-of-hearing people by aiding them with secure resources that improve their communication autonomy. By such initiative, it is more inclusive society, giving everyone the opportunity to fully participate and be understood in their preferred language. Apart from advancing technical capabilities, this work focuses on deepening our understanding of how machines can securely learn and interpret the subtle language of signs. This paper introduces a novel CNN architecture designed to improve the accuracy and effectiveness of automatic sign language interpretation and to address key security concerns. Our approach ensures the strength and dependability of the systems by combining hand landmark data with advanced CNN capabilities, safeguarding users from possible risks and increasing their trust in the technology. When a data is taken in form of images and hand landmark are introduces, security point of concern is resolved. Security for data usage and model evaluation is maintained.

CONTRIBUTION OF THIS PAPER

Development of a secure, advanced, and efficient sign language interpretation system focused on deep learning and computer vision technology are few of the contributions of this paper. Firstly, Integration of Hand Landmark Features: To enhance the accuracy and security of sign detection and interpretation, inclusion of hand landmark features ensures more reliable results and reduces error margins. Secondly, Novel Convolutional Neural Network (CNN) Architecture: CNN architecture incorporates hand landmark features as input to improve performance and fortifies the system against security threats by ensuring robust feature extraction and interpretation. This is a notable second contribution from this paper. Thirdly, Comprehensive Performance and Security Analysis: The performance analysis of the proposed hand landmark - enhanced CNN model using a benchmark dataset is the third ns the most important contribution of this paper. This evaluation maintains the data inserted in alignment with model's security and reliability in variety of state of art solutions. This proves the volatile range of the model against potential security risks.

Following this, the next section discusses relevant studies and systems developed for this problem domain. Proposed model is detailed in Section 3 and Experimentation is detailed in section 4. The study's conclusion will be inferred based on the performance and security analysis system.

RELATED WORK

This section focuses on identifying the recent studies or systems developed for sign language interpretation or detection using deep learning, computer vision, or any advanced artificial intelligence technologies to understand the current state. Further, this section also highlights these systems' problems and emphasizes the requirements of this work. Deep learning-based sign language recognition using hand gestures was a very efficient approach used by researchers with different learning architectures. For example, Abdulhussein and Raheem [7] proposed a CNN-based architecture and edge detection method for accuracy improvement and trained their model with the ASL dataset. The other method proposed by Abiyev et al. [8] was hybrid and used a Single Shot Multi Box Detection (SSD) with a support vector classifier for better accuracy. Transfer learning was introduced along with fine-tuning the CNN model for accuracy improvement in Arabic sign language [9]. The improving CNN models and diversity in feature extraction in CNN is a result of better efficiency and availability of multiple CNN designs which resulted in better disease detection specifically for tomato disease detection [40]. Further, the model based on Deep CNN, named E-model, was proposed [10]. This E model consists of various convolutional, max pool, and dense layers and achieves accuracy comparable to the AlexNet models. To enhance the performance of sign language recognition, a novel hand gesture along with the full background space image was used with 3DCNN layers, and these features were fused further before the classification layer [11]. Another 3DCNN model was trained with large-scale data for further improvements [12]. Wadhawan & Kumar [13] proposed 50 different CNN models and evaluated their performance to find the best and optimal model for accurate detection. Furthermore, VGG-11 and VGG-16 were also proposed and tested on two different datasets [14]. However, the accuracy of deep Lenet-5 architecture is better than that of VGG-16 [15], as tested on the ASL dataset. With advancements in deep learning architectures, a feature-based neural translation model was proposed [16]. Apart from these, inception v3 [17], robust CNN [18], LSTM and GRU [19], CNN[20] [21], various one sided shot detection technique using 2D CNN decomposition with singular value(SVD), and LSTM [22], a hybrid Neural Machine Translation (NMT) with Media Pipe and with Dynamic Generative Adversarial Network (GAN) model [23], media pipe with feed-forward neural network [24], CNN, two bidirectional LSTM layers and connectionist temporal classification (CTC) [25], BPN with Histogram Oriented Gradient (HOG) features [26], Mask Region based Convolution Neural Network and shark smell optimization with soft margin support vector machine [27], ResNet with self-attention [28], MobileNetv2 [29], CNN and transformer [30], multistage graph convolution with attention and residual connection [31], YOLO-NAS-S model [32], sep-TCN, deep learning layer, and a channel attention module graph convolution, [31], Graph and General deep-learning network [33], Vision Transformers [34], CNN, bi-directional LSTM, and connectionist temporal classification [35], and CNN with SIFT [36] are in brain tumor [42][45], are some of the recent developments in the sign language detection and recognition system.

Several deep learning models in alignment with real world problem statements are designed and invented in recent years. However, improvement is still in the scope of this problem domain for signature recognition again CNN with diverse feature extraction is used to improve the accuracy and ignore the human presence [41] . With this aim, this work proposes a hand landmark-enhanced CNN model that improves recognition and interpretation.

METHODS

This study presents the latest developments in a CNN model for sign language interpretation that utilizes hand landmark information. By concentrating on the intricate geometry and dynamics of hand gestures—essential for effectively communicating meaning in sign languages—this novel method dramatically increases the accuracy and dependability of sign language identification. Using hand landmark features as inputs, the model anticipates particular motions or signals within a sign language, which helps the deaf and hard-of-hearing people communicate more effectively. Figure 1 presents the overall proposed architecture of this proposed system, and the details of all the steps are discussed as follows:

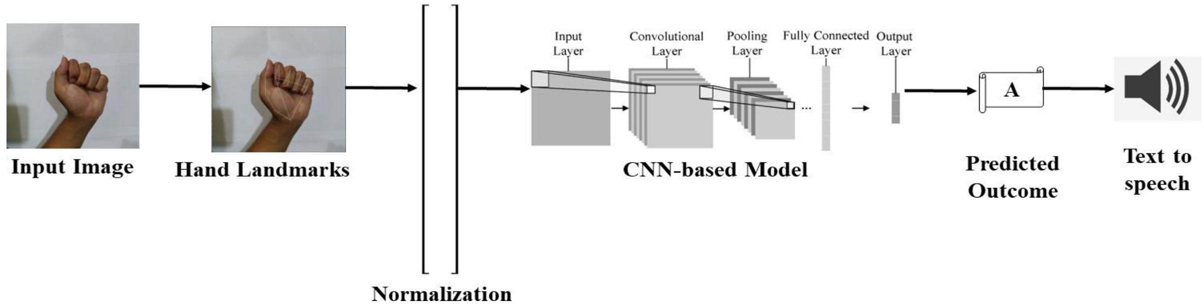

Figure 1. Proposed System Architecture

The process starts by taking a hand image as input for the system where the ASL and SIBI dataset [37] is used. Further, different vital points are extracted and normalized using the hand landmark detection approach. These normalized features are given to the CNN model as input, passing through different layers and finally predicting the outcome based on the classification outcome. Further, the average user verifies the detected sign using the text-to-speech conversion method.

1. Hand Landmark detection

Using MediaPipe, an open-source framework created by Google for creating multimodal applied machine learning pipelines, the hand landmarks are extracted as the cornerstone of this proposed technique [38]. Modern high-fidelity hand and finger tracking technologies are provided by MediaPipe, allowing for the detection of crucial hand locations necessary for deciphering sign language. With just one image, MediaPipe's Hand Tracking solution uses machine learning to infer a hand's 21 3D landmarks [39]. These landmarks identify numerous hand anatomical points, including the fingers, joints, and palms. First, a hand is detected in the image frame. Next, a machine learning model is used to anticipate the landmark. The result is a list of coordinates representing the locations of each landmark in the image. Extracting these hand landmarks can capture a complete depiction of hand gestures, including the form, orientation, and motions of the hand and fingers.

2. Proposed CNN model

Using MediaPipe, Proposed model extracts hand landmarks. The next step is to feed these landmarks into the proposed CNN model. An improved CNN architecture was created, as shown in Figure 2, to efficiently process the landmark data and extract patterns and characteristics indicative of certain sign language motions. The following layers define the proposed CNN:

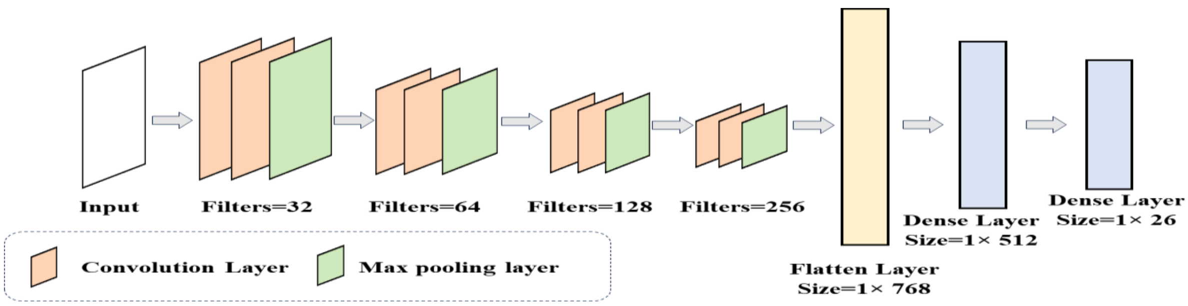

Figure 2. Proposed CNN architecture

3. Input Layer: 21 hand landmark normalized coordinates are accepted as input by the input layer. Each

landmark's x, y, and z coordinates capture the hand motions' three-dimensional structure.

3.1 Convolutional Layers: Several convolutional layers use hand landmarks to learn feature representations. These layers use filters to identify the spatial connections between landmarks, which is necessary to distinguish between different types of sign motions. The figure also defines the number of filters used at different convolution levels.
3.2 Functions of Activation: By introducing non-linearity, the Rectified Linear Unit (ReLU) activation function enables the model to recognize intricate patterns in the data. This activation function and the convolution layers remove the negative features.
3.3 Pooling Layer: Pooling layers are used to lower the dimensionality of the feature maps, reducing computing costs and preventing overfitting. Usually, max pooling is used to save the most essential characteristics.
3.4 Flatten Layer: A flattening layer transforms the multi-dimensional feature maps into a one-dimensional vector before moving on to the dense layers. Connecting the convolutional portion of the network with the thick layers requires this change.
3.5 Dense Layers: Dense layers, which replace fully connected layers, integrate the high-level information learned throughout the convolutional stages. This integration is essential for the task of gesture prediction. These layers are necessary for matching specific sign language movements or signals to the retrieved landmark attributes.
3.6 Output Layer: Using a softmax activation function, the design ends with an output layer. To classify the inputs, this layer uses pre-established categories that correlate to different sign language signals or gestures.

With the addition of hand landmark characteristics, this improved CNN model creates a new standard for automatic sign language interpretation. The model offers excellent accuracy and efficiency in recognizing a wide variety of sign language movements by utilizing exact information from hand motions, significantly improving communication options for the deaf and hard-of-hearing groups.

EXPERIMENTATION AND RESULT ANALYSIS

The performance of the proposed hand landmark-enhanced CNN is analyzed based on its detection, loss, and accuracy. The experiments use a window-based system with graphic capabilities, and Python is used as a programming language to implement the model. The details of the dataset, parameters, and experimentation are delineated in this section.

1. Dataset details

Two different datasets were used to train and assess the model: the Sistem Isyarat Bahasa Indonesia (SIBI) Alphabets Sign Language dataset (which was confirmed by LEMLITBANG SIBI) and the American Sign Language (ASL) dataset. Images from the ASL collection include a variety of hand forms and lively movements, covering a broad spectrum of ASL signals. The SIBI dataset provides a wide range of gestures for model training by including signals unique to the Indonesian sign language system. Data augmentation techniques were used on both datasets to increase the number and variety of the training data and strengthen the model's capacity for generalization across several sign languages. The sample images of the SIBI dataset for each English alphabet sign are presented in fig 3. Training Parameters

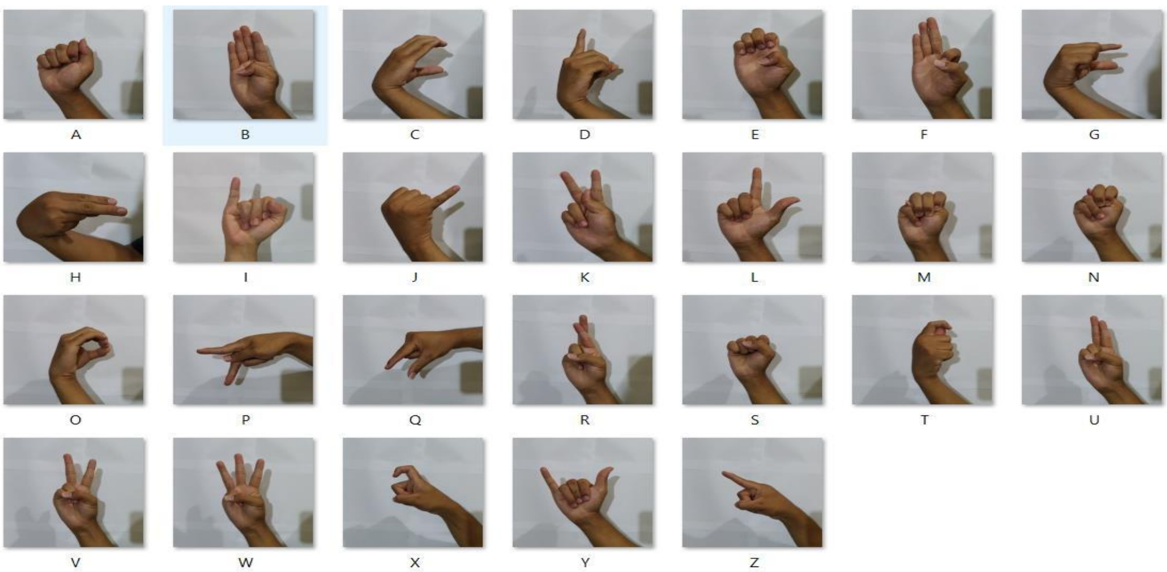

Figure 3. Sample dataset images

The proposed model is trained with the abovementioned datasets and uses hand-landmark features. Hence, its input size depends on the number of features extracted using the feature extraction algorithm. Further, the Adam optimizer is well known for working well with noisy data and sparse gradients.
Based on initial studies, a learning rate of 0.001 was selected to guarantee a reasonable compromise between training stability and fast convergence. A batch size of 32 and 100 epochs of training was used since this combination was proven efficient in reducing training time and increasing model accuracy.

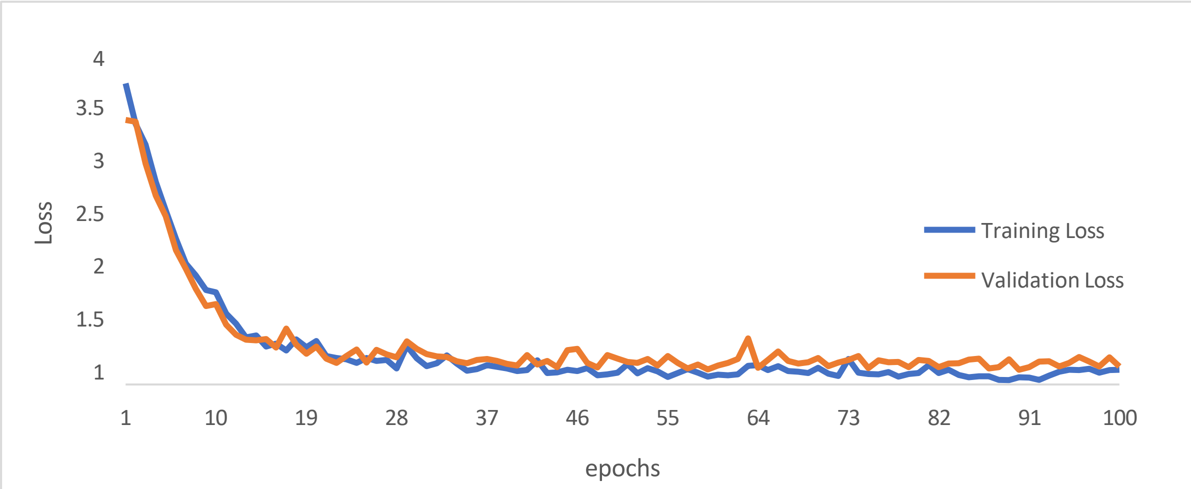

Figure 4. Loss Computation

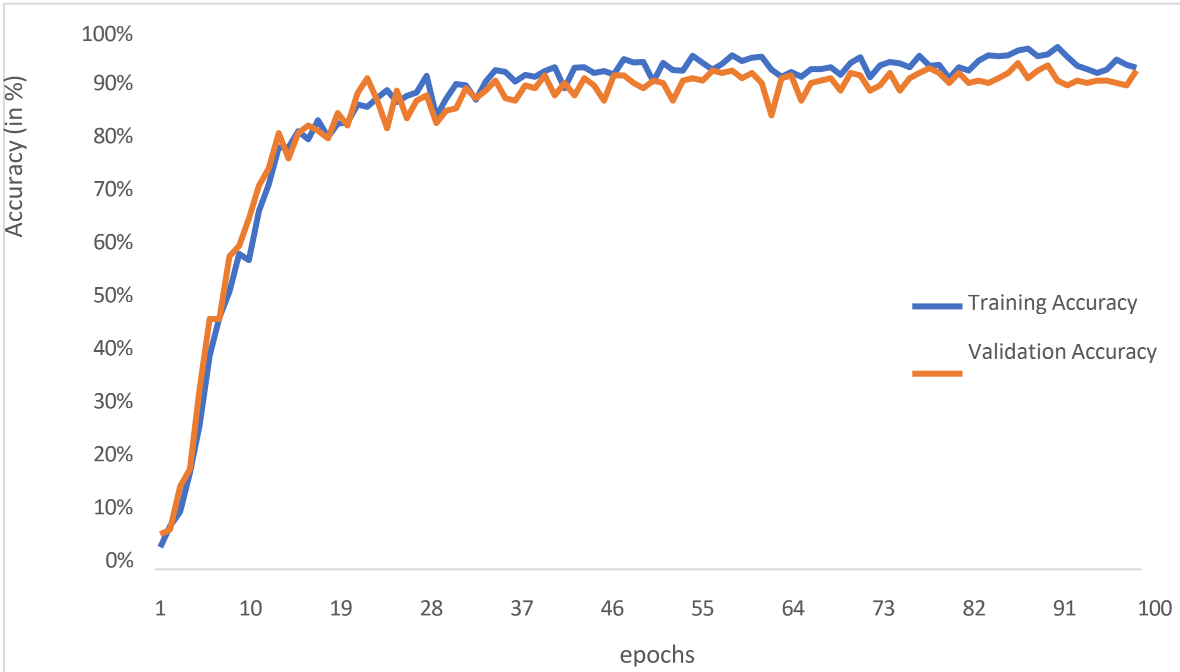

Figure 5. Accuracy Computation

The loss during the training phase is computed and presented in Figure 4. A proposed CNN model for sign language interpretation's reported loss data across 100 epochs demonstrates a notable initial drop in training and validation losses, suggesting that the model is successfully learning from the input data. The training loss moves from 3.684 to 0.1799, and the validation loss from 3.2374 to 0.2198. These substantial drops in losses indicate the model's increasing predictive power over time. However, variations in validation loss, such as the peak from epoch 17 to 0.6852, may point to overfitting as they indicate periods of decreased generalization to new data.

RESULTS

Interestingly, the model's highest generalization performance occurs at epochs 59 and 90, when validation losses are at their lowest (0.1841 and 0.1799). The validation loss fluctuates at epoch 59, suggesting the model may have reached its learning limit without overfitting. Stabilizing loss values after training indicates that the model has learned as much as it can with this configuration. These indicators show the need to balance applying fresh data and learning from training sets. They also point out possible areas for model enhancement, such as regularization or architectural modifications.

1. Accuracy Computation

Accuracy is an essential measure for any deep learning system. The sign language interpretation model's training and validation accuracy findings across 100 epochs demonstrate a notable and steady improvement in performance, as shown in Figure 5. The model appears to be learning from the data initially since the training and validation accuracy are initially low, at 4% and 7%, respectively. There is a noticeable rising trend in the training and validation accuracies as epochs go by. It suggests that the model is successfully picking up new information and applying it to the validation data from the training set. The improvement in the early epochs demonstrates the model's capacity to identify the simpler patterns in the sign language data quickly. The training accuracy of the model exceeds 40%, and the validation accuracy approaches 60% around epochs 6 to 8, indicating the model's increasing proficiency in sign language interpretation. This pattern persists, with both accuracies increasing as training continues, most notably at epoch 21 when training accuracy was 88% and validation accuracy was 90%. Nevertheless, there are times when validation accuracy decreases compared to earlier epochs, such as epochs 14 and 24. These variations may represent the fluctuating complexity of the validation set samples or lead to overfitting at specific moments when the model performs better on the training data than on unobserved validation data. The model reaches its best accuracy in the latter epochs, epoch 88 onwards, with training accuracy peaking at 98% and validation accuracy peaking at 95%. These findings imply that the model has developed into a skilled sign language interpreter capable of extrapolating its comprehension to novel, unobserved data.

2. Detection Rate

After the training and validation stages are finished, a different test dataset is used to evaluate the model's performance. This dataset is essential because it includes sign pictures the model hasn't seen in training or validation. It makes it an unbiased measure of the model's capacity to generalize and interpret sign language correctly in real-world situations. The testing stage is crucial to assessing the model's effectiveness since it ensures it has learned the fundamental patterns and characteristics required for sign language interpretation, not just how to memorize the training data.

The model shows a 99.5% detection rate in the testing phase, which yields surprisingly favorable results. This outstanding accomplishment highlights the model's accuracy and dependability in correctly deciphering the provided sign pictures. With a 99.5% detection rate, the model effectively recognizes and understands each sign presented from the test dataset, ensuring that no sign is misclassified. Such a high degree of accuracy throughout the testing phase demonstrates a successful model that has grasped the nuances of sign language. Some of the detection findings are shown in Figure 6 to highlight further the model's capabilities and the efficacy of its sign language interpretation. These outcomes provide a visual validation of the model's correctness by highlighting particular examples of the model's interpretation of specific indications. These examples demonstrate how well the model captures and identifies the complex motions and combinations that makeup sign language. In addition to verifying the model's functionality, the figure sheds light on the interpretive process of the model. It shows how cutting-edge machine learning methods may help deaf and hard-of-hearing people communicate more effectively.

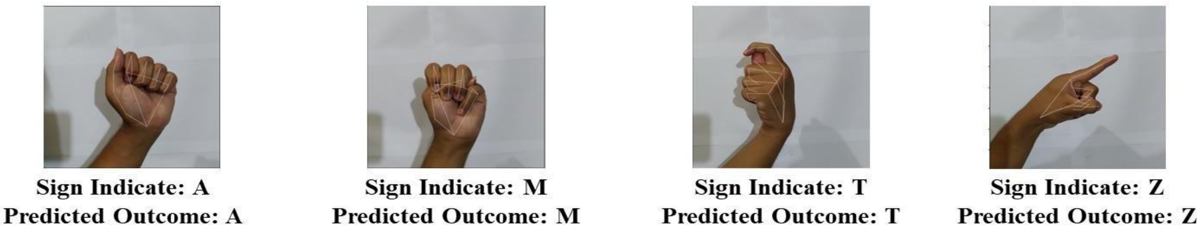

Figure 6. Detection Results

3. Comparison with state-of-the-art (SOTA) methods

The detection accuracy of the proposed model is compared with the recently developed SOTA methods and is presented in the table below. The comparison of several models for the interpretation of sign language demonstrates various techniques and technology, as shown in Table 1, ranging from using SIFT features in conventional CNNs to more sophisticated methods such as Vision Transformers and the suggested method that attained 100% accuracy. Test accuracies show how well each model performs and how many methodologies have evolved to capture the subtleties of sign language using visual inputs. Notably, the combination of real-time tracking, attention processes, feature extraction techniques, and sequence modeling shows a variety of strategies for improving interpretation accuracy.

Table 1. Comparison with SOTA methods

Models	Test Accuracy
CNN with SIFT [36]	98.74%
Vision Transformer [34]	99.3%
LeNet-5 [15]	87.5%
Inception v3[4]	93%
LSTM and GRU[19]	97%
Mediapipe with feed-forward neural network[24]	99%
ResNet with self-attention [28]	98.2%
Proposed Method	99.5%

The flawless score of the proposed model is a noteworthy advancement that underscores the possibility of creative approaches to improve accessibility to sign language and close the gap for the deaf and hard-of-hearing populations.

DISCUSSION

In this paper, the model suggests a significant improvement in automation in sign language interpretation using machine learning and artificial intelligence with computer vision with a strong focus on security. The model during review with comparison to various models from traditional CNN models and SIFTs etc with proposed model clearly explains a remarkable change in achieving this accuracy. This also shows potential of accuracy with ensuring the data integrity and privacy of sign language. The diverse methodologies with ongoing improvements in accuracy and security illustration in dynamic nature in this field with research and commitment for such individuals. The proposed method not only sets a new benchmark for performance but also showcases how innovative solutions can effectively address the challenges of accurately interpreting sign language. By integrating robust security measures, our approach significantly enhances the reliability and trustworthiness of sign language interpretation systems, protecting against potential threats such as data tampering and unauthorized access. This study advances the technological frontier, has profound social implications, and paves the way for future developments in integrated and inclusive communication technologies. By improving accessibility and reducing communication barriers, our work contributes to a more inclusive society, ensuring that everyone can participate fully and be understood in their preferred mode of communication. The given approach doesn't only set new performance standards but also shows how innovative and intellectual problem-solving approach can effectively address the barriers of a perfect sign language understanding. Our approach significantly boosts the dependability as well as the credibility of sign language interpretation systems by including security features which safeguards against potential vulnerabilities like data abuse and unlawful access. This research offers new opportunities for seamless and inclusive systems of communication, extends the boundaries of technology, and has huge social impacts. Our work contributes to a more inclusive society by reducing communication barriers and enhancing accessibility, allowing everyone to fully participate and be understood in their own way of communication.

REFRENCES

[1] H. Kim, "What constitutes professional communication in aviation: Is language proficiency enough for testing purposes?" *Lang. Test.*, vol. 35, no. 3, pp. 403–426, Jul. 2018, doi: 10.1177/0265532218758127.

[2] D. Bragg *et al.*, "The FATE Landscape of Sign Language AI Datasets," *ACM Trans. Access. Comput.*, vol. 14, no. 2, pp. 1–45, Jun. 2021, doi: 10.1145/3436996.

[3] D. Bragg *et al.*, "Sign Language Recognition, Generation, and Translation," in *the 21st International ACM SIGACCESS Conference on Computers and Accessibility*, New York, NY, USA: ACM, Oct. 2019, pp. 16–31. doi: 10.1145/3308561.3353774.

[4] L. Zheng, B. Liang, and A. Jiang, "Recent Advances of Deep Learning for Sign Language Recognition," in *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, IEEE, Nov. 2017, pp. 1–7. doi: 10.1109/DICTA.2017.8227483.

[5] M. Werngren-Elgström, Å. Brandt, and S. Iwarsson, "Everyday activities and social contacts among older deaf sign language users: Relationships to health and well-being," *Occup. Ther. Int.*, vol. 13, no. 4, pp. 207–223, Dec. 2006, doi: 10.1002/oti.218.

[6] R. McKee, J. Iseli, and A. Murray, "Sign language interpreting in the Pacific: A snapshot of progress in raising the participation of deaf people," *J. New Zeal. Pacific Stud.*, vol. 7, no. 2, pp. 185–196, Oct. 2019, doi: 10.1386/nzps_00005_1.

[7] A. Abdulhussein and F. Raheem, "Hand Gesture Recognition of Static Letters American Sign Language (ASL) Using Deep Learning," *Eng. Technol. J.*, vol. 38, no. 6, pp. 926–937, Jun. 2020, doi:

10.30684/etj.v38i6A.533.
[8] R. H. Abiyev, M. Arslan, and J. B. Idoko, "Sign Language Translation Using Deep Convolutional Neural Networks," *KSII Trans. Internet Inf. Syst.*, vol. 14, no. 2, Feb. 2020, doi: 10.3837/tiis.2020.02.009.
[9] Y. Saleh and G. F. Issa, "Arabic sign language recognition through deep neural networks fine-tuning," *Int. J. online Biomed. Eng.*, vol. 16, no. 5, pp. 71–83, 2020, doi: 10.3991/IJOE.V16I05.13087.

[10] Y. Pratama, E. Marbun, Y. Parapat, and A. Manullang, "Deep convolutional neural network for hand sign language recognition using model E," *Bull. Electr. Eng. Informatics*, vol. 9, no. 5, pp. 1873–1881, Oct. 2020, doi: 10.11591/eei.v9i5.2027.
[11] M. Al-Hammadi *et al.*, "Deep Learning-Based Approach for Sign Language Gesture Recognition with Efficient Hand Gesture Representation," *IEEE Access*, vol. 8, pp. 192527–192542, 2020, doi: 10.1109/ACCESS.2020.3032140.
[12] S. Sharma and K. Kumar, "ASL-3DCNN: American sign language recognition technique using 3-D convolutional neural networks," *Multimed. Tools Appl.*, vol. 80, no. 17, pp. 26319–26331, Jul. 2021, doi: 10.1007/s11042-021-10768-5.
[13] A. Wadhawan and P. Kumar, "Deep learning-based sign language recognition system for static signs," *Neural Comput. Appl.*, vol. 32, no. 12, pp. 7957–7968, Jun. 2020, doi: 10.1007/s00521-019-04691-y.

[14] S. Sharma and S. Singh, "Vision-based hand gesture recognition using deep learning for the interpretation of sign language," *Expert Syst. Appl.*, vol. 182, p. 115657, Nov. 2021, doi: 10.1016/j.eswa.2021.115657.
[15] S. Chavan, X. Yu, and J. Saniie, "Convolutional Neural Network Hand Gesture Recognition for American Sign Language," in *2021 IEEE International Conference on Electro Information Technology (EIT)*, IEEE, May 2021, pp. 188–192. doi: 10.1109/EIT51626.2021.9491897.
[16] T. Ananthanarayana *et al.*, "Deep Learning Methods for Sign Language Translation," *ACM Trans. Access. Comput.*, vol. 14, no. 4, pp. 1–30, Dec. 2021, doi: 10.1145/3477498.
[17] M. Varsha and C. S. Nair, "Indian Sign Language Gesture Recognition Using Deep Convolutional Neural Network," in *2021 8th International Conference on Smart Computing and Communications (ICSCC)*, IEEE, Jul. 2021, pp. 193–197. doi: 10.1109/ICSCC51209.2021.9528246.
[18] S. Sharma and S. Singh, "Recognition of Indian Sign Language (ISL) Using Deep Learning Model," *Wirel. Pers. Commun.*, vol. 123, no. 1, pp. 671–692, Mar. 2022, doi: 10.1007/s11277-021-09152-1.

[19] D. Kothadiya, C. Bhatt, K. Sapariya, K. Patel, A.-B. Gil-González, and J. M. Corchado, "Deepsign: Sign Language Detection and Recognition Using Deep Learning," *Electronics*, vol. 11, no. 11, p. 1780, Jun. 2022, doi: 10.3390/electronics11111780.
[20] A. KASAPBAŞI, A. E. A. ELBUSHRA, O. AL-HARDANEE, and A. YILMAZ, "DeepASLR: A CNN based human computer interface for American Sign Language recognition for hearing-impaired individuals," *Comput. Methods Programs Biomed. Updat.*, vol. 2, p. 100048, 2022, doi: 10.1016/j.cmpbup.2021.100048.
[21] A. El Zaar, N. Benaya, and A. El Allati, "Sign Language Recognition: High Performance Deep Learning Approach Applyied to Multiple Sign Languages," *E3S Web Conf.*, vol. 351, p. 01065, May 2022, doi: 10.1051/e3sconf/202235101065.
[22] R. Rastgoo, K. Kiani, and S. Escalera, "Real-time isolated hand sign language recognition using deep networks and SVD," *J. Ambient Intell. Humans. Comput.*, vol. 13, no. 1, pp. 591–611, Jan. 2022, doi: 10.1007/s12652-021-02920-8.
[23] B. Natarajan *et al.*, "Development of an End-to-End Deep Learning Framework for Sign Language Recognition, Translation, and Video Generation," *IEEE Access*, vol. 10, pp. 104358–104374, 2022, doi: 10.1109/ACCESS.2022.3210543.
[24] J. Bora, S. Dehingia, A. Boruah, A. A. Chetia, and D. Gogoi, "Real-time Assamese Sign Language Recognition using MediaPipe and Deep Learning," *Procedia Comput. Sci.*, vol. 218, pp. 1384–1393, 2023, doi: 10.1016/j.procs.2023.01.117.
[25] S. Sharma, R. Gupta, and A. Kumar, "Continuous sign language recognition using isolated signs data and deep transfer learning," *J. Ambient Intell. Humans. Comput.*, vol. 14, no. 3, pp. 1531–1542, Mar. 2023, doi: 10.1007/s12652-021-03418-z.
[26] R. Sreemathy, M. Turuk, I. Kulkarni, and S. Khurana, "Sign language recognition using artificial intelligence," *Educ. Inf. Technol.*, vol. 28, no. 5, pp. 5259–5278, May 2023, doi: 10.1007/s10639-022-11391-z.
[27] M. M. Alnfiai, "Deep Learning-Based Sign Language Recognition for Hearing and Speaking Impaired People," *Intell. Autom. Soft Comput.*, vol. 36, no. 2, pp. 1653–1669, 2023, doi: 10.32604/iasc.2023.033577.
[28] D. R. Kothadiya, C. M. Bhatt, A. Rehman, F. S. Alamri, and T. Saba, "SignExplainer: An Explainable AI-Enabled Framework for Sign Language Recognition with Ensemble Learning," *IEEE Access*, vol. 11, pp. 47410–47419, 2023, doi: 10.1109/ACCESS.2023.3274851.

[29] M. Daniel Nareshkumar and B. Jaison, "A Light-Weight Deep Learning-Based Architecture for Sign Language Classification," *Intell. Autom. Soft Comput.*, vol. 35, no. 3, pp. 3501–3515, 2023, doi: 10.32604/iasc.2023.027848.

[30] J. Shin *et al.*, "Korean Sign Language Recognition Using Transformer-Based Deep Neural Network," *Appl. Sci.*, vol. 13, no. 5, p. 3029, Feb. 2023, doi: 10.3390/app13053029.

[31] A. S. M. Miah, M. A. M. Hasan, S. Nishimura, and J. Shin, "Sign Language Recognition Using Graph and General Deep Neural Network Based on Large Scale Dataset," *IEEE Access*, vol. 12, pp. 34553– 34569, 2024, doi: 10.1109/ACCESS.2024.3372425.

[32] K. Priya and B. J. Sandesh, "Developing an Offline and Real-Time Indian Sign Language Recognition System with Machine Learning and Deep Learning," *SN Comput. Sci.*, vol. 5, no. 3, p. 273, Feb. 2024, doi: 10.1007/s42979-023-02482-w.

[33] A. S. M. Miah, M. A. M. Hasan, Y. Tomioka, and J. Shin, "Hand Gesture Recognition for Multi-Culture Sign Language Using Graph and General Deep Learning Network," *IEEE Open J. Comput. Soc.*, pp. 1– 12, 2024, doi: 10.1109/OJCS.2024.3370971.

[34] A. F. Alnabih and A. Y. Maghari, "Arabic sign language letters recognition using Vision Transformer," *Multimed. Tools Appl.*, Mar. 2024, doi: 10.1007/s11042-024-18681-3.

[35] Y. Gu, H. Oku, and M. Todoh, "American Sign Language Recognition and Translation Using Perception Neuron Wearable Inertial Motion Capture System," *Sensors*, vol. 24, no. 2, p. 453, Jan. 2024, doi: 10.3390/s24020453.

[36] S. Arooj, S. Altaf, S. Ahmad, H. Mahmoud, and A. S. N. Mohamed, "Enhancing sign language recognition using CNN and SIFT: A case study on Pakistan sign language," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 36, no. 2, p. 101934, Feb. 2024, doi: 10.1016/j.jksuci.2024.101934.

[37] I. G. B. H. Widhinugraha and E. Rakun, "Indonesian Language Sign System (SIBI) Recognition Using Threshold Conditional Random Fields," in *Proceedings of the 2019 8th International Conference on Computing and Pattern Recognition*, New York, NY, USA: ACM, Oct. 2019, pp. 380–384. doi: 10.1145/3373509.3373591.

[38] C. Lugaresi *et al.*, "MediaPipe: A Framework for Building Perception Pipelines," Jun. 2019, [Online]. Available: http://arxiv.org/abs/1906.08172

[39] Y. Lin, X. Jiao, and L. Zhao, "Detection of 3D Human Posture Based on Improved Mediapipe," *J. Comput. Commun.*, vol. 11, no. 02, pp. 102–121, 2023, doi: 10.4236/jcc.2023.112008.

[40] Mamta Gehlot, Rakesh Kumar Saxena, "Tomato-Village": a dataset for end-to-end tomato disease detection in a real-world environment ", Multimedia Systems, https://doi.org/10.1007/s00530-023-01158-y, Springer, 21August 2023.

[41] Deepak Moud, Rakesh Kumar Saxena (2023), Development of signature recognition system using VGG16", Journal of Discrete Mathematical Sciences and Cryptography, April 2023. (Scopus/ Q1 Indexed)

[42] Sunil Gupta; Rakesh Saxena; Ankit Bansal; Kamal Saluja; Amit Vajpayee; Shikha, A case study on the classification of brain tumour by deep learning using convolutional neural network, AIP Conference Proceedings 2782, 020027 (2023) doi: https://doi.org/10.1063/5.0154417.

[43] Sunil Gupta; Mohammad Shahid; Ankur Goyal; Rakesh Kumar Saxena; Kamal Saluja, Black Hole Detection and Prevention Using Digital Signature and SEP in MANET, IEEE Xplore, Volume, Year 2022, Pages 1-5.

[44] Shikha Thakur, Rakesh Kumar Saxena, Analysis of the Functional Relationship between Electrocardiographic Signal and Simultaneously Acquired Respiratory Signals, International Journal of Image, Graphics and Signal Processing, Volume 7(9), Year 2015, Pages 33-40.

[45] Rakesh Kumar Saxena, N. Sumitra, Brain Tumor Classification Using Back Propagation Neural Network, International Journal of Image, Graphics and Signal Processing, Volume 5(2), Year 2013, Pages 45-50.