Real Time Indian Sign Language Recognition usingConvolutional Neural Network

Snehal Hon, Manpreet Sidhu, Sandesh Marathe, Tushar A. Rane
dept. of Information Technology
PICT, Pune
Pune, India

***Abstract*—Through sign language recognition software in particular, technology has made great strides in recent years in bridging the communication gap for those with hearing and speech impairments. To interpret sign language gestures, these systems make use of cutting-edge methodologies like deep learning, computer vision, and neural networks. For instance, a study on Indian Sign Language employed real-time processing and grid-based characteristics to recognise different hand movements without the usage of external hardware. Another method for identifying and categorising American Sign Language motions made use of the HSV colour scheme and Convolutional Neural Networks (CNN). These developments mark significant steps forward in advancing hearing-impaired people's communication, encouraging social integration, and cultivating a deeper comprehension of sign language in a variety of circumstances.**

***Index Terms*—Convolutional Neural Network (CNN), Computer Vision, MediaPipe Holistic, Recurrent Neural Networks, SLR (Sign Language Recognition), SLT (Sign Language Translation).**

I. Introduction

A growing number of people are experiencing hearing problems, particularly in densely populated nations like In- dia, which highlights the need for creative ways to remove obstacles to communication. For the deaf and hard-of-hearing community, technology, in particular developments in speech-to-text and sign language recognition systems, stands out as a crucial tool. This study explores the enormous potentialof these technologies, highlighting their critical function in enhancing access to critical services and bridging communication barriers. The study examines issues including accuracy and accessibility, highlighting the need for more research and development to produce inclusive communication tools designed for users of Indian Sign Language (ISL).

Due to the general public's poor comprehension of In- dian Sign Language (ISL), deaf-mute people confront special social difficulties. Although useful, traditional sign language interpretation techniques can have substantial downsides such high costs and discomfort. New techniques and algorithms for effectively, precisely, and affordably detecting Indian sign language alphabets and words are being made possible by developments in machine learning and deep learning technologies. An extensive solution for improved communication and social integration that is specifically suited for the Indian setting is provided by this study, which carefully examines related work, describes our technique, gives outcomes, and considers limitations and prospects.

In our commitment to bridging the communication gap for individuals with hearing impairments, we have developeda robust sign language recognition system rooted in Indian Sign Language (ISL). Understanding the crucial role languageplays in human interaction, particularly for those with hearing disabilities, our system is designed to facilitate seamless communication using ISL. Sign language, essential for the deaf community, presents unique challenges in conveying its nuances to non-sign language users. To address this, our focus has been on the static recognition of ISL gestures, harnessing the potential of advanced technologies such as Convolutional Neural Networks (CNN). Our innovative approach eliminates the discomfort associated with traditional glove-based methods and enhances portability, achieving exceptional accuracy rates above 90%. By specifically tailoring our system to ISL, we aim to enhance precision and effectiveness, ensuring inclusive communication experiences for everyone involved.

II. Methodologies

The methodology portion of this survey study describes the systematic approach we used to evaluate and contextualize the large corpus of earlier work on sign language recognition using CNN. The methodological complexity of this section, which supports our study, directs the reader. We'll go over the procedures we followed to select the research articles we wanted to analyse, the criteria we used to decide which studies should be included, how we gathered the data, howwe carried out the analysis, and how we assessed the qualityof the studies. By presenting our methodology's guiding principles, we hope to demonstrate its dependability and rigour. This will open the door for a systematic and insightful analysis of sign language recognition research utilising CNN.

In [1] creating datasets from real-time sign language gestures captured via MediaPipe Holistic and OpenCV, the researchers employed CNN models with three convolutional layers for high-precision static sign recognition, outperforming LSTM models. Their LSTM model, featuring three LSTM layers, excelled in recognizing dynamic gestures, achieving

a flawless 100% accuracy for gesture-based signs and an F1 score of 0.8557 for static signs. This research underscored the versatility of LSTM models in gesture recognition and the CNN models' effectiveness in static sign language recognition.

In [2], The study focused on optimizing sign language gesture recognition through careful preprocessing , resizing images to 128x128 pixels, and employing the Histogram Oriented Gradient (HOG) method for feature extraction, ensuring stability across various backgrounds. Feature selection using the LASSO algorithm reduced features from 8100 to 8000, enhancing accuracy and conserving memory. Their use of the squeezenet architecture led to an impressive 97.5% accuracy, emphasizing the efficiency of their approach in sign language gesture recognition.

The methodology in [3] an exploration of Indian Sign Language (ISL) gesture and hand pose recognition, researchers processed a dataset of 24,624 images using advanced techniques like HOG-based face detection and linear SVM classifiers. Their approach ensured precise hand extraction and movement tracking via OpenCV. By employing a Grid-based fragmentation method and feature reduction through PCA and t-SNE, their system achieved impressive results: 99.714% accuracy for ISL hand poses and an average accuracy of 97.2% for recognizing Wrong Gestures (WR) using K-Nearest Neighbours (k-NN) algorithms. The integration of Hidden Markov Models (HMMs), trained with 12 HMM chains via the Baum-Welch algorithm, highlighted the effectiveness of their approach in the nuanced realm of ISL gesture and hand pose recognition.

In [4], the realm of American Sign Language (ASL) gesture recognition, a dataset of 2000 curated images featuring 10 static alphabet signs was meticulously processed and divided for training and testing. Ensuring consistency in lighting and webcam capture, the images underwent transformation into the HSV colorspace for stable background removal. Advanced segmentation techniques enabled precise hand gesture extraction, resulting in binary images resized to

64x64 pixels and generating 4096 features. The model's accuracy in recognizing the 10 alphabetic ASL gestures exceeded 90.0%, reaching a specific accuracy of 98%using Convolutional Neural Networks post normalizing and rescaling.

The methodology in [5] video standardization to 3 seconds with 72 frames and resizing to 64x64 pixels was a focal point. Overcoming data limitations, data augmentation techniques were applied, ensuring consistent transformations. Efficient hand detection algorithms identified hands in video frames, generating augmented datasets. Feature extraction methods included Histogram of Oriented Gradients (HOG) and transfer learning with VGG16. For sequence processing, Long Short Term Memory (LSTM) networks with two stacked layers were optimized using Adam's stochastic gradient function. The methodology integrated frame standardization, data augmentation, advanced hand detection, feature extraction, and LSTM networks, resulting in a model achieving up to 90% accuracy across 20 classes for impactful sign language communication solutions.

In [6], the study centered on Indian Sign Language detection employing a unique dataset and two preprocessing techniques: skin segmentation and running averages. Data preprocessing methods included HSV color space transformation and grayscale conversion with Gaussian filtering. Feature extraction utilized BOVW (Bag Of Visual Words) and SURF (Speeded Up Robust Features) descriptors, followed by K-Means clustering. Classification was carried out using SVM and CNN models, achieving 99.14% accuracy for test data with SVM and over 94% accuracy on the training dataset and over 99% accuracy in testing with CNN.

The methodology of[7] the study employed a self- created American Sign Language Alphabet (ASLA) dataset containing 104,000 static images. Utilizing CNN with multiple layers involving convolution, pooling, flattening, and fully connected layers, images sized 64x64x3 were processed. Feature extraction encompassed 3 convolutional layers, resulting in a training accuracy of 99.38% and a lossof 0.0250. When compared to other datasets, although their dataset showed slightly lower accuracy, it was attributed to its larger size and the incorporation of newly generated features and conditions.

In this methodology[8], ASL (American Sign Language) alphabet recognition was performed using Mediapipe and LSTM. The approach involved utilizing a dataset created by the Mediapipe open-source library, capturing hand gesturesfor 30 frames and storing them as NumPy arrays for model training. A total of 93,600 NumPy arrays were generated, indexed, and stored. The extracted features were then passed to the LSTM Model for recognition. The network architecture included six layers: three LSTM layers and three dense layers, structured as a sequential model for efficient training. This approach resulted in high precision, recall, and F1-score for most alphabets, achieving an impressive accuracy, Micro Average, and Weighted Average of 0.99 for their custom dataset.

In this methodology[9], a system based on an LSTM-GRU model was developed by leveraging the IISL2020 dataset, which consisted of videos featuring 11 words in Indian Sign Language (ISL) from 16 different individuals. Feature extraction was performed using InceptionResNetV2, and overfitting was mitigated by incorporating a dropout layer. The final output was obtained through a softmax function. Six different combinations of LSTM were explored, and the models were tested across various datasets, demonstrating an impressive accuracy rate of 95%. Notably, their dataset achieved the highest accuracy, reaching 97% compared to others.

In this study[10], a replacement system utilizing a three-layer CNN was proposed, utilizing a dataset comprising 87,000 images of American Sign Language (ASL). The model achieved a perfect 100% accuracy during training and a commendable 91% accuracy during testing. Data Augmentation techniques such as Rotation and Cropping were employed, facilitated by a data generator. Background elimination was performed using the HSV color space model, and images were segmented. Feature extraction involved extracting highlights by scaling the images to 64 pixels. The outcomes showed a 78% accuracy for alphabet recognitionand an 80% accuracy for recognizing common gestures.

The methodology of [11] used the DETR (Detection Transformer) approach, which is based on the Vision Transformer and aims to improve the accuracy of sign language recognition at the current state-of-the-art level. The DETR method described in this research accurately recognises sign language from digital videos using a novel deep learning model ResNet152 + FPN (i.e., Feature Pyramid Network), which is based on Detection Transformer. The recently constructed net ResNet152 + FPN can achieve upto 1.70 percent increase in detection accuracy on the sign language test dataset compared to the traditional Detection Transformer models. ResNet50 is replaced with ResNet152as the backbone. It adds additional convolutional layers and works to improve the feature map. Additionally, the suggested approach resulted in an overall accuracy of 96.45%.ResNet50 is replaced with ResNet152 as the backbone. Itadds additional convolutional layers and works to improvethe feature map.

In [12], The ability of a machine to identify human behaviour is known as human activity recognition, or HAR. HAR, a well-known application of state-of-the-art machine learning and artificial intelligence techniques, uses computer vision to understand the semantic meanings of a range of human actions. This research’s supervised learning method uses data from real human movements to distinguish between different human occupations. The primary barrier to adopting HAR is overcoming the difficulties caused by the cyclostationary nature of the activity signals. This research proposes a two-channel Convolutional Neural Network (CNN) based HAR classification model, learning from thefrequency and power characteristics of the collected human action signals. The model was tested using the UCI HARdataset, and the classification accuracy of the results was 95.25 percent. Further research on the recognition of human activity based on biological signals can be assisted by this tactic.

[13] The major goal of this work is to create a deep learning-based application that translates sign language into text and enables communication between signers and non-signers. They developed their own CNN (Convolutional Neural Network) to recognise the sign from a video frame. The MNIST dataset is used.The CNN model is trained using the MNIST ASL dataset. A data set including 27455 training samples and 784 attributes is used to train the model. The model was trained with cross entropy ADAMto minimise loss as much as possible. Ten epochs are used to train the various model on a batch size of 128. Training was conducted at a learning rate of 0.001 with no degradation. There are 7172 samples in the validation dataset, and the model’s stated validation accuracy is higher than 93%.

[14] This aims to train a real-time sign language detector for the given dataset by transfer learning utilising a TensorFlow object identification API. To collect data, images from a camera are taken using OpenCV and Python. The dataset was created for Indian Sign Language, in which the signs correspond to the English alphabet. The data gathering method is used to create the dataset. Transfer learning hasbeen used to train it on the generated dataset, which comprises650 photos overall—25 images for each letter. At 10,000 steps, the total loss experienced during the final training phase was 0.25; the losses for localization, classification, and regularisation were 0.18, 0.13, and 0.10, respectively.At step 9900, the lowest loss of 0.17 was experienced. At 10,000 steps, the total loss experienced during the final training phase was 0.25, with losses in localization comingin at 0.18, classification coming in at 0.13, and regularisation coming in at 0.10. At step 9900, the lowest loss of 0.17 was experienced.Their method has obtained an average confidencerate of 85.45%, despite the modest size of the dataset.

In [15], analysis of various machine learning methods were done. KNN methods using different extraction techniques lead to an accuracy between 63% (Euclidean distance) and 99.23% (KNN using Dynamic Time Warping (DWT)). ANN methods lead to an accuracy between 62% for ANN with backpropagation, and 99% using ANN with HOG. SVM methodologies had highest accuracy of 99.82% using CNN based on AlexNet and VGG16 models. Fuzzy logic led to 100% with 19 rules and 97.5% with ten rules. Ensemble learning methods when used with SVM lead to 98.7%. Among the lesser popular machine learning algorithms, Graph Convolutional Network (GCN) lead to 98.08% accuracy.

The research in [16] offers a real-time ISLR system that uses the Darknet-53 convolutional neural network in conjunction with the YOLOv3 Model. Real-time testing of the system has been conducted using 16 distinct indicators for photos and 7 signs for videos. The datasets for sign language were labelled in YOLO format using the suggested model. For static sign language recognition, the webcam records images in sign language, while for dynamic sign language recognition, videos are recorded. For both static and dynamic indicators, we were able to reach accuracy levels of 95.7% and 93.1%, respectively. The experimental findings demonstrate that the suggested system can efficiently and in real time, with less computational time, recognise both static and dynamic ISL indicators.

III. Analysis

The synthesis and analysis section of a survey paper plays a crucial role in distilling the surveyed literature. This section provides a succinct summary of existing research, emphasizing prevalent trends, critically evaluating the strengths and weaknesses of different methods, and offering valuable insights for future research directions.

In this section, an in-depth analysis of research conducted over the past two decades on the topic will be presented. This review will focus on identifying prevalent trends, patterns, and recurring themes observed in the surveyed research papers. The methodologies employed will be critically evaluated, emphasizing their strengths and weaknesses. Additionally, the interconnections between the surveyed research papers will be explored, and the practical implications of the findingsand methods in real-world scenarios will be discussed. Furthermore, potential avenues for future exploration in the field will be considered.

The studies discussed in the realm of sign language recognition collectively demonstrate the depth and diversity of methodologies employed in this field. A hybrid approach, combining CNN and LSTM models, as seen in [1], highlights the flexibility needed for recognizing both static and dynamic sign language. This fusion proved effective, showcasing the potential of diverse neural network architectures in addressing the complexities of sign language gestures [1].

Moreover, papers like [2] emphasize the crucial roleof preprocessing and feature extraction techniques. Image resizing and the application of Histogram Oriented Gradient (HOG) methods offer stability across diverse backgrounds and lighting conditions [2]. The utilization of the LASSOalgorithm for feature selection and the choice of efficient pre- trained models like Squeezenet underscore the significance of meticulous method selection for achieving high accuracy rates [2].

Additionally, the studies in [3], [4], and [5] delve into the nuances of specific sign languages, showcasing the depth of understanding required for accurate recognition. Integration of Hidden Markov Models (HMMs) in [3] highlights the need to capture variations in gestures performed by different signers. Advanced segmentation techniques and color space transformations in [4] provide a solid foundation for precise gesture extraction, leading to specific accuracies of 98%in recognizing ASL gestures. The comprehensive approachin [5], involving standardized videos, diverse pre-trained models for feature extraction, and the incorporation of Long Short Term Memory (LSTM) networks, showcases the multidimensionality of efforts required for effective sign language recognition, resulting in impressive accuracies of up to 90% across various classes [5]. These studies collectively underscore the interdisciplinary nature of sign language recognition, where innovative techniques from computer vision, machine learning, and deep learning converge to bridge the communication gap for the speech-impairedcommunity.

In the landscape of sign language recognition, a variety of innovative methodologies have surfaced, each tailored to the complexities of specific sign languages and gestures. Paper 6 introduced a meticulous approach, combining skin segmentation and running averages with traditional techniques like grayscale conversion and Gaussian filtering [6]. Employing BOVW with SURF descriptors and clustering via K-Means, this study achieved remarkable accuracy. Additionally, the integration of SVM and CNN models demonstrated the efficacy of a comprehensive preprocessing and classification strategy [6].[7] focuses on the American Sign Language Alphabet (ASLA) dataset, leveraging a CNN architecture enriched with convolution, pooling, flattening, and fully connected layers [7]. Despite slightly lower accuracy during training, the study's emphasis on dataset size, including 104,000 static images, highlighted the strength derived from a large dataset and distinctive feature generation [7].

An innovative ASL alphabet recognition system, incorporating Mediapipe and LSTM technologies [8].Their model, comprising three LSTM layers and three dense layers, achieved exceptional precision, recall, and F1-score [8].In [9] pioneered the use of LSTM-GRU models for Indian Sign Language(ISL) recognition, achieving an impressive accuracy rate of 97% [9]. Their custom dataset (IISL2020) and the employment of InceptionResNetV2 for feature extraction were pivotal. Through various LSTM combinations and the strategic use of dropout layers, this study emphasized the significance of tailored approaches in achieving high accuracy rates for specific sign languages [9].Meticulous approach resulted in a training accuracy of 100% and a testing accuracy of 91% [10]. These studies collectively emphasize the nuanced nature of sign language recognition, showcasing the importance of tailored preprocessing techniques, diverse feature extraction methods, and thoughtful model architectures in achieving accurate and effective sign language interpretation [6][9][10].

One prominent theme in the analyzed papers is the extensive use of convolutional neural networks (CNNs) to achieve accurate recognition. The study in [11] adopted the ResNet152 + FPN model, achieving a noteworthy accuracyof 96.45% by replacing ResNet50 as the backbone. Similarly, in [13], a CNN model trained on the MNIST ASL dataset achieved a validation accuracy surpassing 93%. These findings underscore the effectiveness of CNNs in the domain of sign language recognition.

Furthermore, paper [12] explored the application of CNN in human activity recognition (HAR) and demonstrated the capability of CNNs to distinguish between various human occupations with a high classification accuracy of 95.25%. This highlights the versatility of CNNs not only in sign language recognition but also in broader domains likeHAR. Moreover, the research presented in [14] leveraged transfer learning in conjunction with TensorFlow's object identification API for real-time sign language detection. Despite a relatively small dataset, this approach achieved an average confidence rate of 85.45%, showcasing the promiseof transfer learning in real-time sign language recognition.

The exploration of different machine learning techniques in [15] unveiled the remarkable performance of support vector machine (SVM) methodologies, particularly those using CNNs based on AlexNet and VGG16 models, which reached an impressive accuracy of 99.82%. This underlines the significance of selecting suitable machine learning algorithms for sign language recognition tasks. Lastly, in [16], the authors presented a real-time Indian Sign Language Recognition (ISLR) system employing the Darknet-53 convolutional neural network and YOLOv3 Model. Their system achieved 95.7% accuracy for static sign language recognition and 93.1% for dynamic sign language recognition,demonstrating the viability of these models for real-time ISL recognition tasks

IV. Block Diagrams

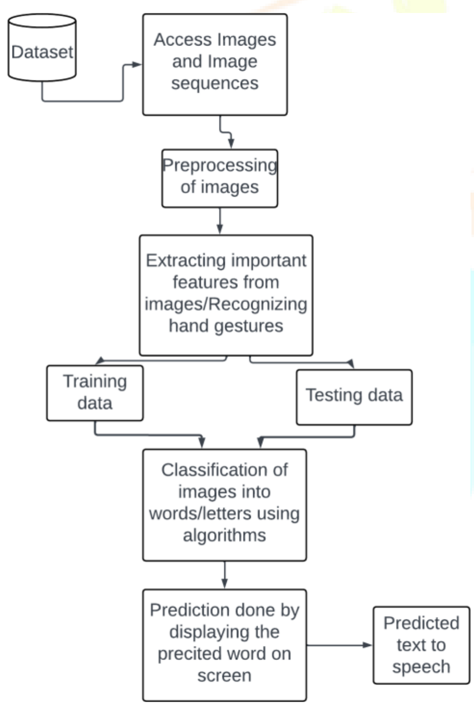

Fig. 1. Data Flow Diagram

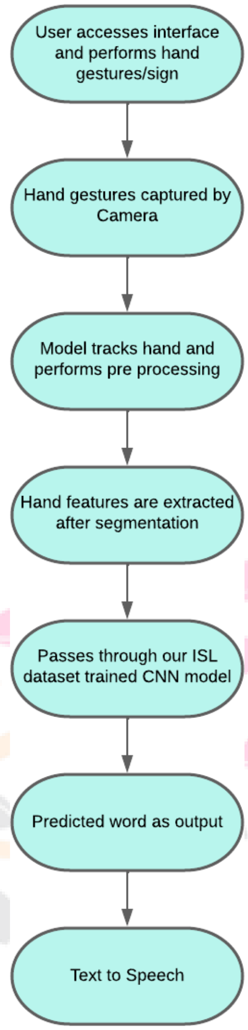

Fig. 2. System Design Diagram

The functional and fundamental flow of our suggested system is depicted in the system design diagram. It displays both what can be done by the user and what the system is capable of doing in real time. To get adequate real-time precision, it is crucial to have both high accuracy and a well-trained model. With the aid of the diagrams, all of this was simply comprehended.

V. Proposed Solution

Based on the analysis of the existing system, In our research on Indian Sign Language (ISL) recognition, we propose a cutting-edge approach by integrating Convolutional Neural Networks (CNN) within a hybrid framework. Thishybrid CNN architecture will enable us to capture both static and dynamic sign language gestures effectively. We can effectively capture alphabets as well as words.

The specialised capabilities of many networks can be used by utilising hybrid CNN. For comparison, we will mostly use CNN with LSTM, but we will also explore other approaches where necessary. Our ISL dataset will become more accurate as a result. We will be able to process both static sign images and sign language gestures with the aid of a hybrid neural network. Due to its hybrid CNN architecture, which integrates many architectures, it improves the performance of neural networks in sign language recognition, which usually involves complicated and varied input.

As there are limited resources for ISL alphabets/words, and to ensure the effectiveness of our approach, we will curate a comprehensive dataset specific to ISL gestures, encompassing diverse signer variations.

Additionally, we emphasize the incorporation of advanced technologies such as Mediapipe for gesture tracking and color space transformations. These techniques will enhancethe precision of gesture extraction and feature engineering, ultimately contributing to the development of a robust and accurate ISL recognition system.

This approach suggests utilising Mediapipe for precise tracking of hands, fingers, and body landmarks, which facilitates more accurate interpretation of sign language gestures because of Mediapipe's robust gesture tracking and extraction capabilities, which are essential in the field of sign language recognition. Its real-time processing capabilities makes it ideal for live sign language exchanges, which is our system's primary goal.

Our suggested approach uses the user's device's camera to recognise sign language in real time. For improved accuracy, hand features are taken from the recorded feed in the same way that they exist in our dataset. Every feature will be compared, mapped, and then classified using our trained and tested dataset. and a text display of the predicted words and alphabets appears on the screen. The last step in promoting diversity will be to turn the final predictions from text tospeech.

Since there has been less study on common ISL wordsand alphabets than ASL, particularly on words, the primary goal of this proposed system is to develop a model for these terms. Through meticulous experimentation and optimization, our methodology aims to significantly improve the accuracy and adaptability of ISL recognition systems, thereby enhancing communication opportunities for the speech-impaired community.

VI. Conclusion

The goal of our survey paper has been to shed lighton the unmet technological needs of the Deaf and Hard of Hearing (DHH) community, with a particular focuson Indian Sign Language recognition. In this pursuit, our study proposes the creation of a Real-Time Indian Sign Language Recognition system using a hybrid CNN approach. Although this system is currently in the conceptual phase, our comprehensive analysis of existing methodologies and their limitations provides in- valuable insights for its future implementation. By advocating for technology that aligns with the unique requirements of the DHH community, our survey paper underscores the significance of socially responsible technological advancements.

By demonstrating the importance of tailored solutions catering to the diverse needs of individuals, especially marginalized communities like the DHH, our research paves the way for a more inclusive technological landscape. This approach emphasizes the necessity of considering social responsibility in technology design, encouraging future researchers and developers to create solutions that prioritize accessibility and inclusivity. Collaborative efforts between researchers, technologists, and the DHH community will be crucial in bringing this vision to fruition.

This advancement has the potential to significantly enhance educational opportunities, improve employment prospects, and elevate the overall quality of life for DHH individuals.By addressing the specific needs of this community, our proposed system aims to bridge the existing technologicalaccessibility gap, ensuring a more inclusive society for all.

References

[1] Goyal, Kaushal. "Indian Sign Language Recognition Using Mediapipe Holistic." arXiv preprint arXiv:2304.10256 (2023).

[2] Mohammedali, Atyaf Hekmat, Hawraa H. Abbas, and Haider Ismael Shahadi. "Real-time sign language recognition system." International Journal of Health Sciences 6 (2022): 10384-10407.

[3] Shenoy, Kartik, et al. "Real-time Indian sign language (ISL) recognition." 2018 9th international conference on computing, communication and networking technologies (ICCCNT). IEEE, 2018.

[4] Murali, Romala Sri Lakshmi, L. D. Ramayya, and V. Anil Santosh. "Sign language recognition system using convolutional neural network and computer vision." (2020).

[5] Puranik, V., Gawande, V., Gujarathi, J., Patani, A., Rane, T.(2022, August). Video-based Sign Language Recognition using Recurrent Neural Networks. In 2022 2nd Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-6). IEEE.

[6] Katoch, Shagun, Varsha Singh, and Uma Shanker Tiwary. "Indian Sign Language recognition system using SURF with SVM and CNN." Array 14 (2022): 100141.

[7] Kasapbas¸i, Ahmed, Ahmed Eltayeb AHMED Elbushra, AL- HARDANEE Omar, and Arif Yilmaz. "DeepASLR: A CNN based human computer interface for American Sign Language recognition for hearing-impaired individuals." Computer Methods and Programs in Biomedicine Update 2 (2022): 100048.

[8] Sundar, B., and T. Bagyammal. "American Sign Language Recognition for Alphabets Using MediaPipe and LSTM." Procedia Computer Science 215 (2022): 642-651.

[9] Kothadiya, Deep, Chintan Bhatt, Krenil Sapariya, Kevin Patel, Ana-Bele´n Gil-Gonza´lez, and Juan M. Corchado. ”Deepsign: Sign language detection and recognition using deep learning.” Electronics 11, no. 11 (2022): 1780.

[10] Piyush Kapoor, Hema N:Sign Language and Common Gesture Using CNN. Electronics (2021):2278-3091

[11] Yu Liu, Parma Nand, Md Akbar Hossain, Minh Nguyen and Wei Qi Yan. Sign language recognition from digital videos using feature pyramid network with detection transforme. Springer(2022):21673–21685

[12] Niloy Sikder, Abu Shamim Mohammad Arif, Md. Sanaullah Chowdhury and Abdullah-Al Nahid. Human Activity Recognition Using Multichannel Convolutional Neural Network. (2019) IEEE

[13] R.S. Sabeenian, S. Sai Bharathwaj and M. Mohamed Aadhil. Sign Language Recognition Using Deep Learning and Computer Vision(2020) ISSN: 1943-023X

[14] Sharvani Srivastava, Amisha Gangwar, Richa Mishra and Sudhakar Singh. Sign Language Recognition System using TensorFlow Object Detection API. Springer 2021

[15] I.A. Adeyanjua, O.O. Bello b and M.A. Adegboye. Machine learning methods for sign language recognition: A critical review and analysis (2021) Elsevier. 2667-3053

[16] Nihashree Sarma, Anjan Kumar Talukdar and Kandarpa Kumar Sarma. Real-Time Indian Sign Language Recognition System using YOLOv3 Model(2021) IEEE