



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: 1 Month of publication: January 2024

DOI: <https://doi.org/10.22214/ijraset.2024.57882>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Survey Paper - Sign Language Detection and Conversion using Deep Learning

Prof. Manjushree Raut¹, Ayush Singh², Khushi Polshettiwar³, Vedika Thakor⁴, Aditya Tibile⁵

Information Technology Department, Smt. Kashibai Navale College of Engineering, Pune

Abstract: Communication for the deaf and mute relies on intricate gestures and visual cues, presenting a challenge for accurate interpretation. This project employs advanced deep learning techniques, utilizing SVM and K-means Clustering algorithms for precise real-time recognition of these expressions from images. Our aim is to bridge the communication gap, empowering the hearing-impaired to engage meaningfully and be a vital part of human interaction. Through this innovative approach, we create an inclusive society where everyone can express themselves effectively, regardless of their hearing abilities.

Keywords: communication, deaf and mute community, gestures, deep learning, SVM, K-means Clustering.

I. INTRODUCTION

In our world, communication is vital, whether it's through touch, voice, or gestures. For people who can't hear or speak, they use hand movements and facial expressions to talk.

Imagine a silent dance where hands and faces tell stories without words. For those who can't speak, these movements are their words. It's like painting pictures with gestures, showing feelings and thoughts without talking.

In this silent world, these gestures aren't just movements; they are a powerful way to share emotions and ideas. It's a reminder of how strong and creative humans can be. Even without words, they find beautiful ways to connect and understand each other.

II. LITERATURE SURVEY

1) *Deep Convolutional Network with Long ShortTerm Memory Layers for Dynamic Gesture Recognition* by Rostyslav Siriak, Inna Skarga-Bandurova, Yehor Boltov at (IEEE – 2019)

In this research, a significant advancement in the field of gesture recognition has been made through the development and implementation of a CNN-LSTM network. The primary problem addressed in this project was the accurate and real-time recognition of hand gestures, crucial for applications in sign language translation and human-computer interaction. The study is grounded in the rapidly evolving field of deep learning and computer vision, with a focus on dynamic gesture recognition. The research begins by acknowledging the success of Deep Neural Networks in various domains, especially in image recognition and classification.

The need for reliable methods in recognizing dynamic gestures is highlighted, emphasizing their applications in contactless interaction, sign language translation, and assisting visually impaired individuals in indoor navigation. Despite the potential, until recently, automatic sign language recognition did not fully utilize the available technology. The proposed solution centers around a CNN-LSTM network, a hybrid architecture incorporating Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) layers. CNNs are widely known for their effectiveness in image-related tasks, while LSTMs are specialized in capturing temporal patterns, making them ideal for analyzing sequential data such as video streams. The architecture of the network involves multiple convolutional layers followed by max-pooling, dense layers, and LSTM layers. Techniques like dropout are utilized to prevent overfitting. The research methodology involves the creation of a labeled dataset of hand gestures and preprocessing steps, including grayscale conversion, segmentation, and simplification. The CNN-LSTM model is trained using this dataset, and extensive experiments are conducted to optimize the hyperparameters and validate the approach. The training process involves constructing a database, data preprocessing, data augmentation, feature identification, and hand gesture recognition. The chosen loss function, categorical cross-entropy, ensures effective training, while the Adam optimizer updates the network weights. The performance evaluation of the model demonstrates exceptional accuracy, on the test dataset. This high accuracy underscores the efficiency of the proposed approach in recognizing hand gestures. The research also emphasizes the stability of the model across varying hand rotation angles and lighting conditions, thanks to the utilization of contour patterns. A normalized confusion matrix illustrates the model's ability to accurately classify different gestures.

In terms of related work, the paper discusses existing gesture recognition methods and models, positioning the proposed CNN-LSTM approach as a significant advancement due to its real-time capabilities, stability, and accuracy.

The utilization of CNN and LSTM layers distinguishes this work from previous solutions, enabling it to effectively handle the complexities of dynamic gestures. The tools and technologies employed in this research include TensorFlow and Keras frameworks, numpy, scikit-learn, and OpenCV libraries. These choices are strategic, aligning with the project's requirements and objectives. The CNN-LSTM model's development and evaluation were performed on a standard computer setup, emphasizing the practicality and accessibility of the proposed solution. In conclusion, this research contributes significantly to the domain of gesture recognition by introducing a robust and efficient CNN-LSTM network. The approach demonstrates superior accuracy and real-time performance, addressing the challenges associated with dynamic gesture recognition. The utilization of established frameworks and libraries ensures the reliability and replicability of the research findings, making it a valuable addition to the field of computer vision and assistive technology.

2) *Sign Language Recognition System Based on Weighted Hidden Markov Model* by Wenwen Yang, Jinxu Tao, Changfeng Xi, Zhongfu Ye at (IEEE -2015)

In this research, a significant advancement in the field of gesture recognition has been made through the development and implementation of a CNN-LSTM network. The primary problem addressed in this project was the accurate and real-time recognition of hand gestures, crucial for applications in sign language translation and human-computer interaction. The study is grounded in the rapidly evolving field of deep learning and computer vision, with a focus on dynamic gesture recognition.

The research begins by acknowledging the success of Deep Neural Networks in various domains, especially in image recognition and classification. The need for reliable methods in recognizing dynamic gestures is highlighted, emphasizing their applications in contactless interaction, sign language translation, and assisting visually impaired individuals in indoor navigation. Despite the potential, until recently, automatic sign language recognition did not fully utilize the available technology. The proposed solution centers around a CNN-LSTM network, a hybrid architecture incorporating Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) layers. CNNs are widely known for their effectiveness in image-related tasks, while LSTMs are specialized in capturing temporal patterns, making them ideal for analyzing sequential data such as video streams. The architecture of the network involves multiple convolutional layers followed by max-pooling, dense layers, and LSTM layers. Techniques like dropout are utilized to prevent overfitting.

The research methodology involves the creation of a labeled dataset of hand gestures and preprocessing steps, including grayscale conversion, segmentation, and simplification. The CNN-LSTM model is trained using this dataset, and extensive experiments are conducted to optimize the hyperparameters and validate the approach. The training process involves constructing a database, data preprocessing, data augmentation, feature identification, and hand gesture recognition. The chosen loss function, categorical cross-entropy, ensures effective training, while the Adam optimizer updates the network weight. The performance evaluation of the model demonstrates exceptional accuracy, achieving 98.46% on the test dataset. This high accuracy underscores the efficiency of the proposed approach in recognizing hand gestures. The research also emphasizes the stability of the model across varying hand rotation angles and lighting conditions, thanks to the utilization of contour patterns. A normalized confusion matrix illustrates the model's ability to accurately classify different gestures.

3) *Indian Sign Language Recognition Using ANN And SVM Classifiers* by Miss. Juhi Ekbote Mrs. Mahasweta Joshi at (IEEE-2018)

The paper delves into the significance of sign language as a primary mode of communication for the deaf and mute community. Sign language offers a structured way of communication, where specific gestures represent words or alphabets. Each region typically has its own sign language, and this research focuses on Indian Sign Language (ISL). While extensive research has been conducted on sign languages like BSL and ASL, ISL has received relatively less attention. The research aims to bridge this gap by developing an automatic recognition system specifically tailored for ISL numerals (0-9). The study uses various techniques for feature extraction, including shape descriptors, SIFT, and HOG. Shape descriptors, such as eccentricity, aspect ratio, compactness, extent, solidity, orientation, and roundness, are employed to characterize the gestures. SIFT, a method for detecting and describing local features in images, is also utilized to extract keypoints from the sign images.

Additionally, the paper discusses the use of HOG, a widely used descriptor for object detection, in recognizing the shape of objects within images. HOG works by analysing gradient orientations in localized portions of an image, making it robust against geometric and photometric variations. The research employs a database of 1000 images, with 100 images representing each numeral sign. The images are pre-processed and segmented, and various features are extracted using the aforementioned techniques. The extracted features are then fed into classification algorithms, including ANN and SVM, for sign recognition.

The results showcase the effectiveness of different combinations of feature extraction techniques and classifiers. The study reveals that a combination of HOG and ANN yields an exceptional accuracy of in recognizing ISL numerals. In conclusion, the research provides a comprehensive overview of the challenges in recognizing ISL gestures and proposes an innovative solution using advanced feature extraction techniques and classifiers. The high accuracy achieved demonstrates the system's efficacy, making it a significant contribution to the field of sign language recognition.

4) *Live Action And Sign Language Recognition Using Neural Network by Mrs.Indumathy P, M. Tech 1, Ms. Nithyalakshmi J 2, Ms. Monisha P 3, Ms. Mythreyee M 4 at (ICICC-2023)*

This research project addresses the crucial need for effective communication tools for the speech and hearing-impaired community. Sign language, a vital means of nonverbal communication, inspired this study. The project proposes a solution involving live action tracking and recognition of sign language gestures through advanced machine learning techniques. The project delves into the challenges faced by the speech and hearing-impaired population, emphasizing the significant percentage of the global population affected by hearing loss. It explores the intricate nature of sign language, highlighting the importance of body movements and gestures in communication.

A survey of existing solutions reveals diverse approaches. Some studies employ LSTM and GRU models for Indian Sign Language, achieving high accuracy. Other researchers use YOLOv5 and CNN for real-time sign gesture recognition, demonstrating substantial improvements. A study also discusses MobileNet architecture for baby sign language classification. These existing methods provide valuable insights into sign language recognition technologies. The project introduces a novel approach using Deep Learning with LSTM based on TensorFlow and Keras. It outlines the modules of the system, including pattern recognition, dataset collection, model training, and deployment. Utilizing MediaPipe for facial detection and OpenCV for dataset creation, the project emphasizes the importance of accurate data collection and preprocessing. The research incorporates cutting-edge technologies such as TensorFlow, Keras, LSTM, and MediaPipe for developing the sign language recognition system. The utilization of OpenCV and Mediapipe for dataset collection and preprocessing ensures high-quality input for the machine learning models. The project acknowledges limitations related to processing speed, considering the time taken for recognizing actions. The throughput might be slow due to the complexity of gesture recognition and language conversion.

In conclusion, the research project offers a comprehensive exploration of sign language recognition. By employing advanced deep learning techniques and leveraging powerful libraries and frameworks, the proposed system aims to bridge the communication gap for the speech and hearing-impaired community. The study contributes to the field by providing a detailed methodology and utilizing state-of-the-art tools, emphasizing the potential for real-time sign language recognition systems in facilitating inclusive communication.

5) *Sign language Recognition Using Machine Learning Algorithm by Greeshma Pala, Jagruti Bhagwan Jethwani, Satish Shivaji Kumbha, Shruti Dilip Patil at (IEEE - 2021)*

The proposed system employs three algorithms: K-Nearest Neighbours (KNN), Multi-class Support Vector Machine (SVM), and Convolutional Neural Network (CNN) for hand sign recognition. KNN classifies gestures based on nearest neighbours, SVM uses the RBF kernel for optimal hyperplane separation, and CNN utilizes convolutional layers to learn distinctive features from images, providing high accuracy. Data Collection: Gathered around 29,000 images of hand signs using a webcam, capturing gestures performed by individuals. Data Preprocessing: Images were converted to grayscale, resized to 75x75 pixels, and transformed into numerical arrays for uniformity and efficient processing. Dataset Split: The dataset was divided into training (75%) and testing (25%) data for model training and evaluation. Algorithm Selection and Configuration: K-Nearest Neighbours (KNN): Chose a suitable 'k' value (3) and used the Euclidean distance metric for classification. Support Vector Machine (SVM): Utilized the Radial Basis Function (RBF) kernel and tuned the gamma parameter (0.001) for classification. Convolutional Neural Network (CNN): Constructed a CNN model with convolutional layers, ReLU activation, max-pooling, and dense layers. Implemented softmax activation for multiclass classification. Implemented a callback function to prevent overfitting.

Training: Models were trained using the prepared training dataset, optimizing for accuracy and minimizing loss. Evaluation: The accuracy and performance of each algorithm were evaluated using the testing dataset. Confusion matrices were generated to analyze classification results. Speech-to-Text Conversion: Integrated speech recognition using microphones and utilized libraries like pyaudio and pyttsx3 for converting speech to text and vice versa. Displaying Results: Recognized gestures were classified into text form, displayed on the screen, and converted to speech using pyttsx3. Comparison and Analysis: Compared the accuracy of KNN, SVM, and CNN models to determine the most effective algorithm for hand sign recognition.

Future Scope: Explored potential applications, such as aiding communication between speechless individuals and others and facilitating sign language learning for wider accessibility.

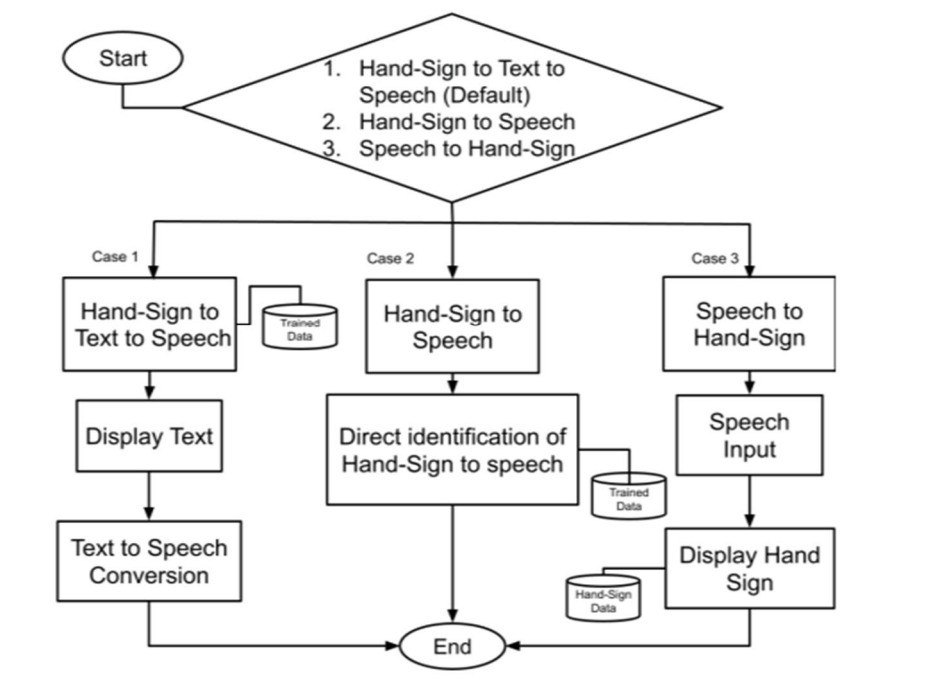


Fig 1. System Flow Diagram

Data Collection: Hand sign gestures performed by individuals in front of a webcam. Speech-to-Text Conversion: Speech input from users captured via a microphone for translation into text.

III. GAP ANALYSIS

Paper Name	Algorithm used	Accuracy of Algorithm	Focus	Data Collection
Deep Convolutional Network with Long Short Term Memory Layers for Dynamic Gesture Recognition	CNN and LSTM	Moderate	SL Recognition	Films of SL
Sign Language Recognition System Based on Weighted Hidden Markov Model	HMM	Moderate to High	Gesture Recognition	Not Specified
Indian Sign Language Recognition Using ANN And SVM Classifiers	SVM & ANN	Moderate	ISL Recognition	Isolated Images
Sign language Recognition Using Machine Learning Algorithm	SVM and random forest	Moderate	ISL Recognition	Dataset

Live Action And sign neural Network Language Recognition Using Neural Network	Neural Network	Moderate	Body Lang Rec.	Live Human Actions
Helping Hearing-Impaired in Emergency Situations: A Deep Learning-Based Approach	3D CNN,VGG-16 with RNN-LSTM, YOLO v5	Moderate	Emergency signs in Indian Sign language(ISL)	Videos for eight different emergency situations
The Comparison of Some Hidden Markov Models for Sign Language Recognition	Modified Hidden Markov model (Gaussian HMM)	Moderate	Argentine Sign Language, edge and skin detection, hand movement	Videos for 10 signs, edge detection, skin detection, hand movement
Design of Sign Language Recognition Using E-CNN	Ensemble of CNN models	Moderate to high	Bridge between the Community and deaf people using image processing	Hand key point library, CNN models, Ensemble method
SIGN LANGUAGE CNNS ALPHABET RECOGNITION USING CONVOLUTION NEURAL NETWORK	CNNs	Moderate to high	Recognition of American Sign Language (ASL) Alphabets	Pre-processed hand gesture images from the MNIST dataset

IV. CONCLUSION

- 1) This project strives to bridge the communication gap between hearing-impaired and hearing individuals. Our deep learning-based approach, employing Convolutional Neural Networks, ensures accurate recognition of symbolic expressions from images.
- 2) By offering real-time capabilities, we aim to create a more inclusive and connected society, allowing everyone, regardless of hearing ability, to communicate effectively and express themselves in a meaningful way.

REFERENCES

- [1] Deep Convolutional Network with Long ShortTerm Memory Layers for Dynamic Gesture Recognition by Rostyslav Siriak, Inna Skarga-Bandurova, Yehor Boltov at (IEEE – 2019)
- [2] Sign Language Recognition System Based on Weighted Hidden Markov Model by Wenwen Yang, Jinxu Tao, Changfeng Xi, Zhongfu Ye at (IEEE -2015)
- [3] Indian Sign Language Recognition Using ANN And SVM Classifiers by Miss. Juhi Ekbote Mrs. Mahasweta Joshi at (IEEE-2018)
- [4] Sign language Recognition Using Machine Learning Algorithm by [Greeshma Pala](#); [Jagruti Bhagwan Jethwani](#); [Satish Shivaji Kumbhar](#); [Shruti Dilip Patil](#) at (IEEE - 2021)
- [5] Live Action And Sign Language Recognition Using Neural Network by Mrs.Indumathy P, M.Tech1*, Ms.Nithyalakshmi J 2, Ms.Monisha P 3, Ms.Mythreyee M 4 at (ICICC-2023)
- [6] Helping Hearing-Impaired in Emergency Situations: A Deep Learning-Based Approach by QAZI MOHAMMAD AREEB 1, MARYAM1, MOHAMMAD NADEEM 1, ROOBAAE ALROOBAAE 2, AND FAISAL ANWER 1 at 2021 IEEE
- [7] The Comparison of Some Hidden Markov Models for Sign Language Recognition by Suharjito, Herman Gunawan, Narada Thiracitta, Gunawan Witjaksono at 2018 IEEE
- [8] Design of Sign Language Recognition Using E-CNN by 1st Citra Suardi, 2nd Anik NurH andayani, 3rd Rosa Andrie Asm ara, 4th Aji Prasetya Wibawa, 5th Lilis Nur Hayati, 6th Huzain Azis at IEEE Xplore – 2021
- [9] SIGN LANGUAGE ALPHABET RECOGNITION USING CONVOLUTION NEURAL NETWORK by Mayand Kumar, Piyush Gupta, Rahul Kumar Jha, Aman Bhatia, Bickey Kumar Shah at ICICCS 2021
- [10] Real-time Sign Language Recognition using Computer Vision by Ruchi Gajjar, Jinallee Jayeshkumar Raval at IEEE Xplore - 2021



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)