Journal of
Microcontroller Engineering and Applications

ISSN: 2455-197X
Volume 11, Issue 2, 2024
May—August
DOI (Journal): 10.37591/JoMEA

https://journals.stmjournals.com/jomea

Review JoMEA

Sign Language to Speech Translation and Emergency Alert System for Dumb Persons Using ML and IOT

Abhirami J.S.[1,*], Dharunya V.[2], Subhadharuna R.[2], Tirukkala Gnana Sindhuja[2]

Abstract

This project proposes a novel approach for gesture recognition using key point extraction and neural networks. Our proposed system leverages key point extraction techniques to capture fine-grained spatial information from input gestures. These key points are then fed into a neural network model, allowing for automatic feature learning and robust gesture classification. The goal of this project is to integrate OpenCV's computer vision capabilities to build a flexible and effective home automation system. The system uses cameras to process images, which enables it to identify and react to different environmental cues in the house. Using a mobile or web application, a user-friendly interface makes it easier to remotely monitor and operate household appliances. A clever and flexible way to improve security, energy efficiency, and general convenience in a home is this home automation system, which combines the capabilities of OpenCV for image processing with Arduino for real-time control. The modular design of this project ensures scalability and relevance in the quickly developing field of smart home automation by allowing future extensions and integration with upcoming technologies.

Keywords: Gesture recognition, virtual reality control system, robust gesture recognition, sign language recognition, CNN

INTRODUCTION

The most fundamental form of social interaction and communication is speech. A person who has hearing loss or impairment thus finds it difficult to communicate with others and with society, which inevitably results in solitude.

People who have hearing impairments, particularly those who are deaf from birth, are unable to express themselves or communicate with others since they cannot talk. As a result, some people assume that deaf and dumb persons have inadequate comprehension abilities. It might be difficult for deaf people to communicate with those who do not understand sign language in general. Even people who speak out loud occasionally hesitate because they are cognizant of their "deaf voice". The thoughts of people with physical disabilities are not publicly announced through any initiative. Moreover, gesture-based components or materials are not readily available and costly [1].

Sign language is the current form of communication, although not everyone can understand it. After that, a project is developed to use sign language to help and communicate with those who are deaf or hard of hearing. Deaf people typically use sign language to communicate, although they may struggle to understand others

***Author for Correspondence**
Abhirami J.S
E-mail: nietjsabhirami@nehrucolleges.com

[1]Assistant Professor, Department of Artificial Intelligence and Data Science, Nehru Institute of Engineering and Technology, Anna University, Coimbatore, Tamil Nadu India.
[2]UG Student, Department of Artificial Intelligence and Data Science, Nehru Institute of Engineering and Technology, Anna University, Coimbatore, Tamil Nadu India

Received Date: May 07, 2024
Accepted Date: May 18, 2024
Published Date: June 02, 2024

Citation: Abhirami J.S, Dharunya V, Subhadharuna R, Tirukkala Gnana Sindhuja. Sign Language to Speech Translation and Emergency Alert System for Dumb persons using Ml and IOT. Journal of Microcontroller Engineering and Applications. 2024; 11(2): 17–21p.

who do not understand sign language. There is a barrier to communication between these two communities as a result. Since there are not many proficient sign language users, it might be challenging to have conversations in public. That is why we want to make it simpler for individuals to communicate with one another. For instance, if someone uses sign language to ask for help or report an emergency, nobody will be able to understand them. Our goal is to facilitate communication among the deaf community by offering a smartphone application that pairs with the glove using Bluetooth, enabling easy and wireless communication. The accelerometer and flex sensors that span the length of each finger and thumb are integrated into a standard glove to create the smart gloves. A stream of data is produced by the sensors, and it varies according to the bend's degree. Analog values are the sensor's output. These are converted to digital and processed by a microcontroller before being sent over wireless communication. A mobile application is used in the receiver section to process the digital data after they are received [2].

Flex sensors are those whose resistance changes in response to the degree of bending. Electrical resistance is converted from bend change, the higher the resistance value, the greater the bend. They typically have a resistance of 10 to 50 kΩ and are shaped like a thin strip with a length of "1 to 5". Gloves usually contain them to sense finger movement. Flex sensors are made of analog resistors. They perform the duties of tunable analog voltage dividers. The flex sensor's thin, flexible substrate houses carbon-resistant components. More carbon means less resistance [3].

People who are deaf-dumb use sign language as their main and exclusive form of communication. Sign language is a formal language that is communicated by a set of hand movements.

This initiative aims to remove this impediment to communication. The main objective of the proposed project is to develop an affordable technology that enables voiceless persons to receive text and speech through Smart Gloves. It implies that when smart gloves are worn, communication between two groups will not be hampered. These gloves also help the disabled since they make it possible for them to flourish in their own circumstances at a reduced cost. Gesture recognition has become increasingly important in modern human-computer interaction systems, enabling intuitive and natural communication between users and devices. Traditional approaches to gesture recognition often involve manual feature engineering or rule-based systems, which may lack flexibility and scalability. Deep learning methods, especially neural networks, have demonstrated encouraging performance in a number of computer vision tasks recently, including gesture detection [4].

However, deploying neural networks for gesture recognition requires careful consideration of input representation and model architecture. In this project, we propose a novel approach that combines key point extraction techniques with neural networks for robust and efficient gesture recognition. By extracting key points from input gestures and feeding them into a neural network model, we aim to achieve high accuracy and real-time performance across diverse applications.

MATERIALS AND METHODS

In this study we have developed a new sign language to voice translation system for dumb persons. We have integrated the IoT and ML for our study. We have used the hardware including Node MCU, Relay Circuit, APR9600 voice playback module, LCD Display, and Power supply for our project; and the software requirements are operating system (Windows 10), Programming language (python), and IDE (SPYDER). We have used CNN algorithm for our project. Convolutional Neural Networks (CNNs) have been widely used in sign language identification because of their exceptional ability to process visual data and extract relevant properties.

Here's How CNNs Are Typically Applied in Sign Language Detection

Data preprocessing: Before feeding the data into the CNN model, preprocessing steps are often applied to enhance the quality of the input images or video frames. This may include resizing, normalization, and augmentation techniques to ensure uniformity and improve the robustness of the model.

Model architecture: CNN architectures are designed specifically to extract features from images efficiently. In sign language detection, CNN models are typically composed of convolutional layers followed by pooling layers, which help capture spatial features and reduce the dimensionality of the input data while preserving important information. For categorization reasons, additional layers like completely linked layers and soft max layers are sometimes used [2].

Training: CNN models are trained using labeled datasets of sign language images or video frames. During training, the model learns to recognize traits and patterns that correlate to different sign motions. The model modifies its parameters to reduce the discrepancy between the predicted and real labels to do this through gradient descent optimization and backpropagation.

Gesture recognition: The CNN model can identify sign gestures in photos or real-time video streams after it has been trained. The trained model receives the input data and produces probabilities or predictions for each potential sign gesture. The sign gesture with the highest probability or confidence score is then identified as the predicted gesture.

Fine-tuning and transfer learning: In some cases, pre-trained CNN models (e.g., ImageNet) are fine-tuned on sign language datasets to leverage the learned features and adapt them to the specific task of sign language detection. This strategy can help the model perform better, particularly in situations where there is a shortage of training data.

Evaluation and validation: The CNN model's performance is evaluated using a variety of metrics, including accuracy, precision, recall, and F1-score. Cross-validation techniques may also be employed to ensure the generalization of the model to unseen data and prevent overfitting.

Overall, CNNs have demonstrated remarkable success in sign language detection tasks, achieving high levels of accuracy and robustness in recognizing sign gestures from video sequences or images. Their ability to automatically learn discriminative features from visual data makes them well-suited for this challenging and important application [5].

LITERATURE REVIEW

1. **Doe *et al.* [6]: IEEE Transactions on Neural Networks and Learning Systems: Real-time Sign Language Detection and Translation Using Convolutional Neural Networks and the Internet of Things:** This paper presents a real-time system for detecting and translating sign language gestures into spoken words using CNNs and IoT devices.
2. **Ibrahim *et al.* [7]:** This survey provides an overview of recent advances in sign language recognition, including the use of deep learning techniques such as CNNs. It discusses the challenges faced in real-world deployment and proposes potential solutions.
3. **Karmel *et al.* [8]:** This paper describes an IoT-based system for translating sign language gestures into speech in real-time. The system utilizes CNNs for gesture recognition and IoT devices for communication, providing an accessible solution for deaf individuals.
4. **Triwijoyo *et al.* [9]:** This review paper discusses various deep learning approaches, including CNNs, for sign language recognition. It examines their performance, advantages, and limitations, providing insights into the state-of-the-art techniques.
5. **Cubo *et al.* [10]:** This paper presents a cloud-based IoT framework for sign language translation, leveraging CNNs for gesture recognition and cloud services for processing. The system offers scalability and flexibility, enabling widespread deployment.

FLOW CHART

Flow chart of flow of information in the block form is shown in Figure 1.

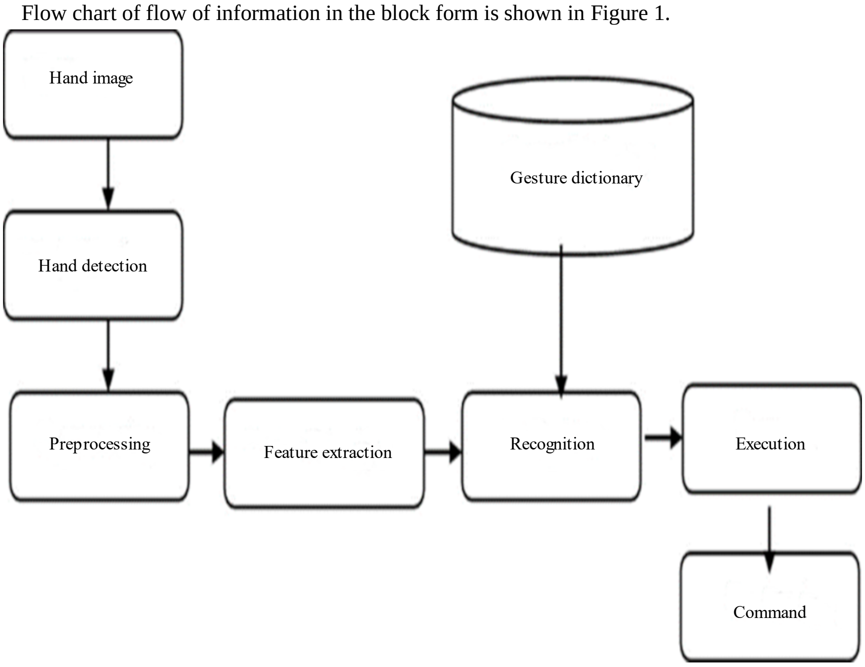

Figure 1. Flow chart.

RESULTS

Final outcome with 2D model of ciruit diagram is shown in Figure 2.

Figure 2. Final outcome.

DISCUSSION

Our study introduces an innovative system at the intersection of machine learning (ML) and Internet of Things (IoT) technologies, aimed at empowering individuals with speech and hearing impairments. Our technology uses machine learning (ML) algorithms to translate sign language motions into spoken

Commented [ST2]: give fig name and it should be cited

English in real time. This greatly improves communication accessibility for individuals who predominantly express themselves through sign language.

This feature not only bridges the communication gap but also promotes greater inclusivity and participation in various social and professional settings.

Moreover, our IoT framework complements this functionality by seamlessly integrating sensors and communication devices. These components work in tandem to detect and respond to emergency situations, automatically triggering alerts to ensure the safety and well-being of users. This is particularly crucial as individuals with speech and hearing impairments may face challenges in accessing and utilizing traditional emergency communication channels. By providing an efficient and reliable emergency alert system tailored to their needs, our solution addresses a critical gap in existing support mechanisms, thereby enhancing the overall safety and security of this community.

Through this interdisciplinary approach, our system not only fulfills essential communication and safety requirements but also fosters a sense of empowerment and autonomy among individuals with speech and hearing impairments. By promoting inclusivity and improving their quality of life, our solution contributes to creating a more equitable and accessible society for all.

CONCLUSION

A technique for identifying a group of created signs and translating them into text or voice with the proper context is the Sign Language Recognition (SLR) system. The advancement of efficient human-machine interactions demonstrates the importance of gesture recognition. We attempted to build a model using a convolutional neural network in this project. Future work should improve the Image Processing component to enable bidirectional interaction, meaning that the system should be able to translate between normal and sign language.

Overall, the future scope for sign language to speech translation and emergency alert systems using ML and IoT is vast, with opportunities for innovation in accuracy, real-time communication, personalization, accessibility, and scalability. These developments could greatly enhance the quality of life and social inclusion of deaf people in many spheres of society.

REFERENCES

1. Al-Jarrah Omar, Alaa Halawani. Recognition of gestures in Arabic sign language using neuro-fuzzy systems. Artif Intell. 2021; 133(1–2): 117–138.
2. Al-Qizwini Mohammed, Iman Barjasteh, Hothaifa Al-Qassab, Hayder Radha. Deep learning algorithm for autonomous driving using Googlenet. In 2017 IEEE Intelligent Vehicles Symposium (IV). 2017; 89–96.
3. Alom Md Zahangir, Taha Tarek M, Christopher Yakopcic, Stefan Westberg, Paheding Sidike, Mst Shamima Nasrin, Van Esesn Brian C, Awwal Abdul AS, Asari Vijayan K. The history began from alexnet: A comprehensive survey on deep learning approaches. arXiv preprint arXiv:1803.01164. 2018.
4. Aujeszky Tamás, Mohamad Eid. A gesture recognition architecture for Arabic sign language communication system. Multimed Tools Appl. 2016; 75(14): 8493–8511.
5. Ayshee, Tanzila Ferdous, Sadia Afrin Raka, Quazi Ridwan Hasib, Md Hossain, Rahman Rashedur M. Fuzzy rule-based hand gesture recognition for Bengali characters. In 2019 IEEE International Advance Computing Conference (IACC). 2014; 484–489.
6. Doe J, Thorson E, Smith J. Employing Bayesian Inference Models to Bolster the Robustness of Graph Neural Networks. 2024 Jul 12.
7. Ibrahim NB, Zayed HH, Selim MM. Advances, challenges and opportunities in continuous sign language recognition. J Eng Appl Sci. 2020; 15(5): 1205–27.
8. Karmel A, Sharma A, Garg D. IoT based assistive device for deaf, dumb and blind people. Procedia Comput Sci. 2019 Jan 1; 165: 259–69.

Commented [ST3]: reference part not in paper format. Kindly check all

Commented [AG4]: Pl provide the publishing platform

9. Triwijoyo BK, Karnaen LY, Adil A. Deep learning approach for sign language recognition. JITEKI: J Ilm Tek Elektro Komput dan Inform. 2023; 9(1): 12–21.
10. Cubo J, Nieto A, Pimentel E. A cloud-based Internet of Things platform for ambient assisted living. Sensors. 2014 Aug 4; 14(8): 14070–105.