SIGN LANGUAGE TRANSLATOR USING MACHINE LEARNING

Prof. Sreenivasa Perumal[1], Hazim Rashid Malik[2], Fabiann P[3], Mohammed Ateeq[4]

[1] Assistant Professor, CSE, AMCEC, Karnataka, India
[2] Student, CSE, AMCEC, J&K, India
[3] Student, CSE, AMCEC, Karnataka, India
[4] Student, CSE, AMCEC, Karnataka, India

ABSTRACT

Sign Language is the only means of communication for the deaf and mute people. But many of the normal people do not know sign language. Thus, it is difficult for the the people who speak in sign language to communicate with those who don't speak the sign language to communicate. This paper extends the previously proposed Convolutional Neural Network (CNN) model for predicting Sign Language using a MobileNetV2-based transfer learning model. The proposed system aims to enable efficient communication for hearing-impaired users by translating sign language gestures into text or speech. The TensorFlow K-NN Image Classifier is used to train the model on the training set. The classifier involves a k-nearest neighbor classifier. The number of classes is determined by the number of unique signs in the dataset, and each class is associated with one sign. The MobileNet model is pre-trained on a large-scale image dataset and fine-tuned on ASL hand sign images to learn discriminative features. After extracting features from the MobileNet model, a KNN classifier is employed for sign language recognition. KNN is a simple yet effective algorithm that assigns a label to an input sample based on the majority class of its k-nearest neighbors in the feature space. In this case, the neighbors correspond to previously seen sign language gestures. The proposed sign language translator system has numerous practical applications, such as aiding individuals with hearing or speech impairments during everyday interactions. Additionally, it can be integrated into educational platforms to support sign language learners and provide inclusive linguistic education opportunities.

Keyword: - *Sign Language translator, K-NN Image Classifier, CNN, MobileNet, TensorFlow, TokBox.*

---

1. INTRODUCTION

The sign language translator aims to convert sign language gestures into spoken or written language, enabling effective communication between the hearing-impaired and hearing communities. This process involves several steps. First, video input of sign language gestures is captured using a camera or other recording devices. Next, the captured video frames are pre-processed to extract relevant features, such as hand shapes, movements, and facial expressions, which are essential for accurate translation.

Once the features are extracted, they are used as input for the machine learning algorithm, specifically the K-Nearest Neighbors (KNN) classifier. The KNN algorithm is a simple yet powerful method that classifies new instances by comparing them to the labelled instances in the training dataset. In the case of sign language translation, the KNN classifier learns from a dataset containing pre-classified sign language gestures along with their corresponding translations.

During the recognition phase, the KNN classifier analyses the extracted features of the input sign language gesture and searches for the closest matching gestures in the training dataset. The KNN algorithm assigns the label of the k nearest neighbors to the input gesture, where k is a user-defined parameter. Finally, the recognized gesture is transformed into spoken or written language, thereby enabling effective communication between individuals who use sign language and those who don't.

2. PROBLEM STATEMENT

Effective and real-time communication between the hearing-impaired community and people who don't understand sign language is a significant issue. Existing tools and systems are either costly, complex, require specialized hardware or processing abilities, or lack real-time interpretation capabilities. Particularly, sign language translation systems based on Machine Learning have various limitations including computational intensity, need for extensive training data, sensitivity to data scale, feature noise and overfitting. There is a pressing need to develop a cost-effective, user-friendly, Realtime sign language translator system using accessible hardware like standard webcams or phone cameras. Implementing such a system would significantly improve communication and social inclusion for the hearing-impaired community Proposed Solution

3. BACKGROUND WORK

Previous SLT models were limited to a particular sign language which meant that a normal person had to learn that language in order to communicate with speech/hearing impaired people. The present model eliminates such hurdles and lets a person train his own model and make his own gestures which are presented in both text and speech form for the person who is most likely to be a speech/hearing impaired individual.

4. OBJECTIVE

i. Develop and implement a robust computer vision model, specifically, to achieve high accuracy in recognizing and interpreting sign language gestures. The model should be trained on diverse datasets to ensure adaptability to various sign languages and gestures.

ii. Incorporate natural language processing algorithms capable of translating recognized sign language gestures into spoken or written language. Ensure the system's versatility by supporting multiple languages, catering to the linguistic diversity of sign language users.

iii. Design the Sign Language Translator (SLT) to provide real-time translation, allowing for instantaneous communication between sign language users and individuals unfamiliar with the language. Optimize the system's speed and efficiency to create a seamless user experience.

iv. Implement an intuitive and user-friendly interface accessible on a variety of devices, including smartphones, tablets, and computers. The interface should not only facilitate the translation process but also provide instant feedback to enhance user engagement and usability. Purpose, Soppe, Applicability

v. Integrate customization features that allow users to personalize the SLT according to their preferences. This includes the ability to adapt the system to specific sign language variations, customize language preferences, and adjust settings to meet individual communication needs.

5. LITERATURE SURVEY

The literature review for this project draws insights from six reference papers, each addressing various aspects of transportation challenges and innovative solutions. The research paper by Zhihao Zhou et al. [1] proposes a sign-to-speech translation system that utilizes stretchable sensor arrays combined with machine learning. Their system aims to translate sign language gestures into spoken language by capturing and interpreting the hand movements using a flexible and wearable sensor array. The authors developed a deep learning model to classify and recognize different hand gestures accurately. The system demonstrated high accuracy and reliability in real-time sign-to-speech translation, contributing to the field of assistive technologies for sign language communication.
The research paper by Muneer Al-Hammadi et al. [2] presents a deep learning-based approach for sign language gesture recognition. Their work focuses on efficient hand gesture representation to improve the recognition accuracy

and reduce computational complexity. The authors propose a feature extraction method that combines optical flow and convolutional neural networks (CNN). The extracted features are used as input to a CNN-based gesture recognition model. The approach achieves notable results in terms of accuracy and speed, showcasing the potential of deep learning techniques in sign language recognition systems.

The research paper by Jesus Suarez and Robin R. Murphy [3] investigates hand gesture recognition using depth images. The authors propose a method that leverages depth data, typically acquired from depth sensors or cameras, for accurate recognition of dynamic hand gestures. Their approach involves preprocessing the depth images, extracting features, and using a support vector machine (SVM) classifier for gesture recognition. The depth-based gesture recognition system demonstrates robustness in real-world scenarios and showcases the potential of depth information for effective hand gesture recognition in various applications.

In the research paper by Saleh Ahmad Khan et al. [4], the authors address the critical need for effective communication between sign language users and non-signers. The proposed system combines Convolutional Neural Networks (CNNs) with customized Region of Interest (ROI) segmentation to achieve accurate sign gesture recognition. By training the model on a custom image dataset containing five sign gestures, the device translates Bangla Sign language into text. Notably, the implementation on a Raspberry Pi ensures portability and user-friendliness. This work contributes significantly to enhancing accessibility for the hearing-impaired community, bridging gaps in communication and promoting inclusivity. By training five sign gestures using a custom image dataset, the device achieves efficient Bangla Sign language to text conversion. The implementation on a Raspberry Pi ensures portability and user-friendliness. Overall, this research contributes to enhancing accessibility for the hearing-impaired community.

In their research titled "Robust Modelling of Static Hand Gestures using Deep Convolutional Network for Sign Language Translation," Singh, Kumar, and Ansari [5] address the critical challenge of recognizing static hand gestures in sign language. The study focuses on leveraging deep learning techniques, specifically convolutional neural networks (CNNs), to automatically learn discriminative features from raw sign images. Unlike traditional methods that rely on manual feature extraction, the proposed CNN-based approach eliminates the need for explicit handcrafted features. Instead, it allows the model to adapt to variations in hand appearance, size, and background clutter. The authors collect a substantial dataset comprising approximately 35,000 sign images representing 100 static signs. This dataset diversity ensures robustness across different users, lighting conditions, and hand orientations. The CNN architecture is designed to capture hierarchical features, enabling it to recognize intricate hand postures. By training on this diverse dataset, the model learns to generalize well, achieving high accuracy in sign recognition tasks. The evaluation likely involves metrics such as accuracy, precision, recall, and F1-score. Singh et al. demonstrate the effectiveness of their system in recognizing static signs, making it a valuable contribution to sign language translation applications. Overall, their work advances the field by combining deep learning with robust modeling techniques, enhancing communication accessibility for hearing-impaired individuals

6. METHODOLOGY

In this project, the Sign Language Translator using machine learning with the KNN (K-Nearest Neighbours) classifier involves the following methodology:

Data Collection:
Gather a dataset of sign language gestures, ensuring diversity and representation of various signs. Each sign should be associated with labelled data indicating the corresponding sign or gesture.

Data Preprocessing:
Clean and preprocess the collected data, including normalization and feature extraction. - Convert images or video frames of signs into a format suitable for machine learning input.

Feature Extraction:
Identify relevant features from the sign language data, such as hand shape, movement, and orientation. Transform raw data into a feature vector that represents the essential Sign Language Translator 23 characteristics of each sign.

Splitting the Dataset:

Divide the dataset into training and testing sets to evaluate the model's performance.

K-Nearest Neighbors (KNN) Algorithm
Choose KNN as the classification algorithm for your Sign Language Translator. KNN classifies new data points based on the majority class of their k- nearest neighbors in the feature space.

Model Training:
Train the KNN classifier using the training dataset. - During training, the algorithm stores the feature vectors and their corresponding labels.

Model Evaluation:
Evaluate the performance of the model on the testing dataset to assess its accuracy and generalization.

Optimization:
Fine-tune the model parameters, such as the value of K in KNN, to improve accuracy and reduce overfitting.

Testing and Deployment:

Test the Sign Language Translator with new sign language inputs to ensure its effectiveness. Deploy the trained model for real-time translation of sign language Sign Language Translator gestures.

7. ARCHITECTURE

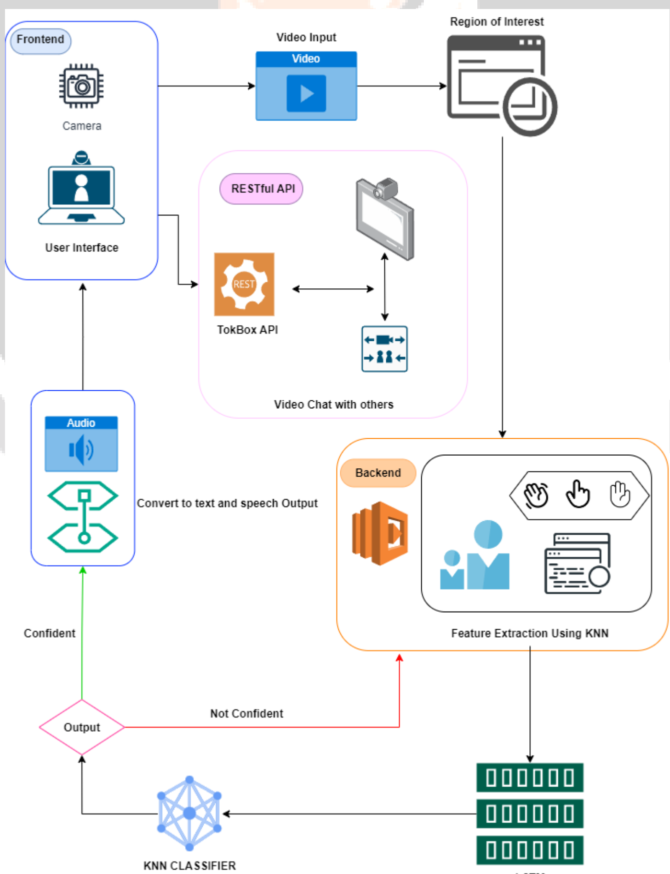

Fig 1: System architecture

The architecture shown in Fig 1 demonstrates how the frontend of our sign language translation webapp interacts with backend to process the given user input and convert/translate it to the corresponding output. The input when processed is categorized based on the confidence level and if it passes a given threshold it is send back to frontend in the form of text and audio. The architecture also shows the implementation of the Video Chat feature using the RESTful API provided by TokBox. With the implementation of TokBox API users are able to connect over a video call with each other. TalkBox uses WebRTC to establish a real time connection between the clients. This real time connection can be between two users or a user and and guest(whose authentication is handled by Firebase). To present real time transcript of the text we used the JavaScript library PeerJS.

8. RESULTS

After successfully implementing the aforementioned application of Sign Language Translator using software tools like NodeJS and Firebase, we have successfully achieved a fully functional real time application. By leveraging innovative technologies like NodeJS, Firebase, Vercel and GitHub, the research has successfully created a comprehensive platform. Through planning, development, and testing, the application overcoming challenges such as limited sign language and learning every possible sign language. This application facilitates speech/hearing impaired people as well as normal people by eliminating the need of an external translator thus proving to be a cost effective and efficient application. With the implementation of TokBox API users are able to connect over a video call with each other. TalkBox uses WebRTC to establish a real time connection between the clients. This real time connection can be between two users or a user and and guest(whose authentication is handled by Firebase).
To present real time transcript of the text we used the JavaScript library PeerJS.

9. CONCLUSIONS

The Sign Language Detector project has offered an innovative solution to bridge the communication gap between the hearing/speech impaired communities and the general population. Utilizing technology, this project has successfully embodied camera sensors, machine learning and software algorithms to interpret sign language accurately. More extended testing and real-world application may be required to fine-tune the system further. Still, initial results depict a major step towards universal communication, promoting inclusiveness and supporting those with hearing or speech impairments. More extended testing and real-world application may be required to fine-tune the system further. Still, initial results depict a major step towards universal communication, promoting inclusiveness and supporting those with hearing or speech impairments.

10. REFERENCES

[1] Zhihao Zhou, Kyle Chen, Xiaoshi Li, Songlin Zhang, Yufen Wu, Yihao Zhou, Keyu Meng, Chenchen Sun, Qiang He, Wenjing Fan, Endong Fan, Zhiwei Lin, Xulong Tan, Weili Deng, Jin Yang and Jun Chen: Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays : Nature Electronics 3, pages571–578 (2020).

[2] Muneer Al-Hammadi, Ghulam Muhammad, Wadood Abdul, Mansour Alsulaiman, Mohammed A. Bencherif, Tareq S. Alrayes, Hassan Mathkour, Mohamed Amine Mekhtiche: Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation: IEEE vol 8 (2020).

[3] Jesus Suarez and Robin R. Murphy: Hand Gesture Recognition with Depth Images: 2012 IEEE RO-MAN: IEEE International Symposium on Robot and Human Interactive Communication. September 9-13, (2012)

[4] S. A. Khan, A. D. Joy, S. M. Asaduzzaman and M. Hossain, "An Efficient Sign Language Translator Device Using Convolutional Neural Network and Customized ROI Segmentation," 2019 2nd International Conference on Communication Engineering and Technology (ICCET), Nagoya, Japan, 2019, pp. 152-156, doi: 10.1109/ICCET.2019.8726895.

[5] D. K. Singh, A. Kumar and M. A. Ansari, "Robust Modelling of Static Hand Gestures using Deep Convolutional Network for Sign Language Translation," 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 2021, pp. 487-492, doi: 10.1109/ICCCIS51004.2021.9397203.