International Journal for Innovative Engineering and Management Research

PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL

www.ijiemr.org

IJIEMR Transactions, online available on 21st Feb 2023. Link

:http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=Issue 03

10.48047/IJIEMR/V12/ISSUE 03/19

Title **SIGNZ - SIGN LANGUAGE RECOGNITION PLATFORM**

Volume 12, ISSUE 03, Pages: 143-147

Paper Authors

Dr. Juliana. R, Veroni Shwetha. D, Raksha. V.G, Elizabeth Subhiksha Victoria

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

SignZ - Sign Language Recognition Platform

Dr. Juliana. R[1], Professor, Department of Information Technology,
Loyola - ICAM College of Engineering and Technology, Chennai.
hodit@licet.ac.in
Veroni Shwetha. D[2], IV B.Tech Department of Information Technology,
Loyola - ICAM College of Engineering and Technology, Chennai.
Raksha. V.G[3], IV B.Tech Department of Information Technology,
Loyola - ICAM College of Engineering and Technology, Chennai.
Elizabeth Subhiksha Victoria[4], IV B.Tech Department of Information Technology,
Loyola - ICAM College of Engineering and Technology, Chennai.

Abstract
This paper proposes a Recognition system for Sign Language using a Convolutional Neural Network (CNN) that recognizes Sign Language digits and alphabets and converts them into speech. This system can help bridge the communication gap between deaf-dumb individuals and the general public. Learning Sign Language can be challenging, and therefore, the proposed system uses an American Sign Language (ASL) dataset to achieve a training accuracy of 99.8%.

Keywords: Convolutional Neural Network (CNN), Python, Google Collab, Google Text to Speech, Open CV, Tensorflow, Sign Language Recognition

Introduction

American Sign Language (ASL) is a commonly used mode of communication by the deaf and hard-of-hearing community. According to recent estimates, there are approximately 1 million ASL users in the US alone, and around 125,000 ASL users in Canada. ASL is also used in other countries, such as Mexico and parts of Central America.

Recognizing sign language poses a significant challenge as signs can differ in motion and appearance, and occur within a continuous flow of gestures. Furthermore, they are often mixed with transitional movements and non-sign patterns, such as out-of-vocabulary signs and epenthesis.

To address this issue, there are two approaches: sensor-based and vision-based. The sensor-based approach involves attaching sensors or inertial measurement units to fingertips and the back of the hand to recognize sign language gestures. However, this approach has limitations as it is not skin-friendly, not cost-efficient, and cannot be worn all the time. In contrast, the vision-based approach has been gaining popularity as it uses machine learning (ML) to recognize sign language.

In our proposed model, we use a novel method that employs Convolutional Neural Networks (CNN) to recognize and classify sign language images. CNN is a popular approach used to train machines for image recognition and classification tasks.

Introduction to the domain

A. Computer Vision

Developing algorithms that allow computers to comprehend and interpret images or videos are the focus of the interdisciplinary field of computer vision. This technology uses deep learning and artificial intelligence techniques to recognize patterns in visual data, allowing computers to analyze, classify, and interpret images in a way that is similar to human perception.

Computer vision has revolutionized many industries, from self-driving cars to medical imaging, and has the potential to transform countless others. In the field of autonomous vehicles, computer vision allows cars to "see" and respond to their

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

surroundings, enabling them to navigate roads and avoid accidents. In medical imaging, computer vision algorithms can assist doctors in diagnosing diseases by analyzing medical images such as X-rays and MRI scans. In manufacturing, computer vision can be used for quality control, ensuring that products meet certain standards before they are shipped. Additionally, computer vision has been used for security purposes, such as identifying faces in surveillance footage and detecting suspicious activity.

Overall, computer vision is a rapidly growing field with endless possibilities for innovation and application. We may anticipate much more intriguing breakthroughs in the future as technology develops and algorithms get more.

B. Tensorflow

Google created the open-source machine learning library known as TensorFlow. Large datasets may be used to create and train machine learning models, which is useful for academics and developers. The library also provides tools for optimizing and deploying machine learning models to various platforms, including mobile devices and the cloud. TensorFlow has become a popular tool for developing applications in various domains, including image and speech recognition, natural language processing, and predictive analytics.

C. OpenCV

The open-source library of software functions known as OpenCV (Open Source Computer Vision) is primarily focused on real-time computer vision. It was created to speed up the deployment of artificial intelligence in commercial products and to provide a common infrastructure for computer vision applications. A variety of features, including object identification, feature detection, and image and video processing, are available with OpenCV. It is compatible with a wide range of different operating systems and programming languages. Applications for OpenCV include robots, driverless cars, and image analysis in the medical field. Both researchers and developers in the computer vision community favour it because of its versatility and usability.

-

Related work

In here, three phase inductive load is In recent years, numerous algorithms have been proposed for recognizing static, dynamic signs, and gestures. Researchers have used various algorithms for sign language recognition (SLR) such as Electrical Impedance Tomography [10], Hidden Markov Model (HMM) [6], 3D Convolutional Neural Networks (3D CNN) [2], and PCANet [12]. For gesture recognition, they have used classifiers such as Deep Neural Network (DNN), Multilayer Perceptron (MLP), Support Vector Machines (SVM), Hidden Markov Models (HMM), and Subspace Gaussian Mixture Model (SGMM) [7].

In the latest study, a system was proposed with 6 layers of CNN to detect digits. The dataset consisted of 20,000 images of 10 BSL digits (0–9), and augmentation performed on a single image generated 50 different variations of images, which provided the highest accuracy of 97.62% [1]. Similarly, a new efficient method for user-independent American fingerspelling recognition based on depth images and PCANet features was proposed, which provided an accuracy of 88.7% [12].

Another study proposed using a nine-layer CNN for 276 sign gestures to accurately recognize sign gestures. It provided an accuracy of an average of 94% to 98.9% in different environments [9]. These different algorithms have demonstrated great accuracy in sign recognition.

Methodology

The proposed model for sign language recognition comprises four distinct modules which work together to interpret hand movements associated with sign language characters in a clear and unambiguous manner. The four modules include data acquisition, pre-processing, feature extraction, and recognition. In the data acquisition module, images and video frames are fed as input to the system. The pre-processing module involves cleaning and filtering the input data to remove noise and unwanted elements. Feature extraction is then carried out, which involves identifying the significant and distinctive features of the input data. Finally, the recognition module makes use of machine learning techniques, particularly Convolutional Neural Networks (CNNs), to classify and predict the sign characters and words displayed in English text, which are also voiced out as words.

To train the model, a novel method using CNN is employed, which is a popular deep learning technique for image recognition and classification tasks. Additionally, the OpenCV library is utilized for real-time operation to recognize the sign characters. This system has significant potential for improving communication between deaf and hearing individuals and bridging the communication gap that exists. It can be used in various real-time applications such as sign language interpretation, education, and accessibility.

A. Data Acquisition

The ASL dataset used in the project was created with a collection of images and video frames. The test dataset was obtained from Kaggle and contained 3000 files for each sign, including ASL digits (0-9) and alphabets (a-z). The dataset included images and videos captured from various angles, positions, backgrounds, and lighting conditions to make it robust and diverse [1].

B. Image Pre-Processing

While training the convolutional neural network to recognize sign language, it was discovered that the variable image sizes in the dataset could not be fed directly into the neural network. In order to address this problem, the images were resized to a standardized proportion of 150 x 150 x 3 pixels using TensorFlow. This rescaling process was necessary for proper image analysis and recognition by the neural network.

Rescaling images is a common pre-processing technique used in deep learning tasks to standardize the image dimensions for effective and efficient training. This technique helps to eliminate any size variation in the input images that may affect the learning and generalization capabilities of the neural network. By rescaling the images, the convolutional neural network can focus on the relevant features of the images and accurately classify them into different categories based on the learned patterns. This process is essential for the successful implementation of the sign language recognition system, as it enables the neural network to accurately classify the input sign language images and predict the corresponding sign character or word.

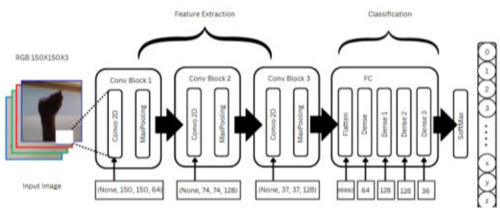
C. Feature Extraction

To recognize hand gestures, the initial step is to isolate the hand region from the rest of the video sequence. The core concept of OpenCV and Mediapipe libraries has been used for feature extraction. The cvzone library acts as an interface for the before-mentioned libraries. Initially, it captures the image in a frame to find the hand and its landmark. It provides information such as the list of landmark points, hand type(left or right) and bounding box. Factors like the width and height of the hand and the distance between the camera and the hand play a role in generating the image. The hand gesture image is created by adding it as an overlay and setting the dimensions of the white image as the threshold size of the image captured.

D. Image Recognition

Python (version 3.10.0) and TensorFlow GPU are used to implement the deep learning framework and learning phase on the server.

The proposed feature extraction is done using the CNN algorithm as shown in the figure given below. It contains three convolutional layers with 64, 128, and 128 filters each, with rectified linear units (ReLu) and MaxPooling following each layer. The proposed model was optimized using the Adam optimizer. The enhanced version of stochastic gradient descent is the Adam optimizer. A learning rate of 0.001 was used in its implementation.



-

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

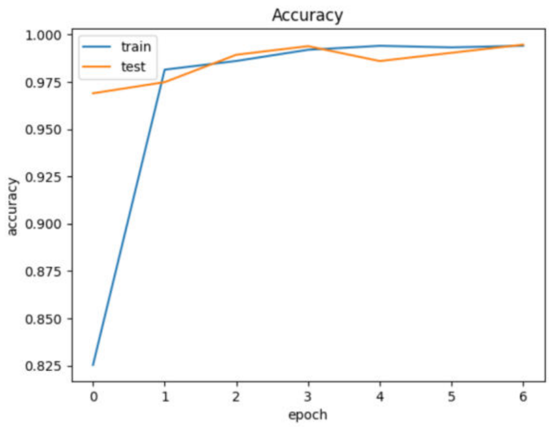
The deployed model is loaded onto a computer or laptop. The signs shown by the user are captured in real-time and scaled into the appropriate size by OpenCV. The model is effective in identifying and predicting words. The suggested system converts ASL letters and numbers into English text [1] before turning them into speech modules.

E. Text To Speech Module

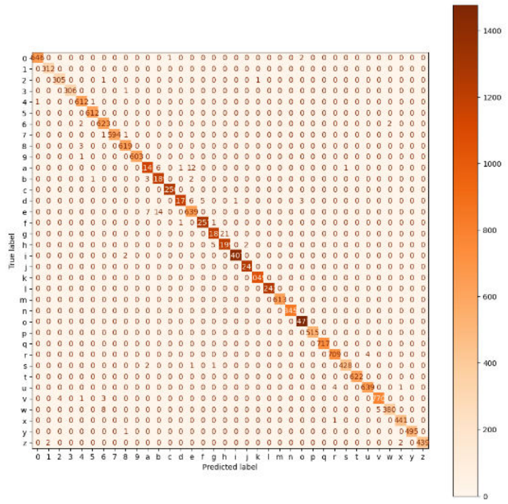
The gestures that are predicted by the system are transformed into voice modules through the utilization of the Google Text to Speech (gTTS) API. This API harnesses the capabilities of Google's advanced machine learning technology, which is capable of converting text into high-quality, natural-sounding speech.

Result

The dataset was split into 80% for training and 20% for testing, and the model was trained using Tensorflow with a batch size of 300. The model was then evaluated using the dataset described in the methodology section, achieving high accuracy with a training accuracy of 99.8% and testing accuracy of 99.46% after 7 epochs of learning using the Adam optimizer. The figure given below compares the accuracy of the proposed CNN model.



The confusion matrix given below indicates that the ASL classification performance was good. However, the model showed improved performance on signs taken from different angles, distances, and illuminations.



Using a webcam, the trained model was able to recognize ASL digits and alphabets in real-time, displaying the predicted sign in text and audio.

Conclusion

In conclusion, the Signz project has demonstrated the potential of using computer vision and machine learning to bridge the communication gap between Deaf and Dumb people and the rest of the world. By leveraging a compact CNN-based architecture, the system is able to accurately and efficiently predict American sign language gestures in real time. The integration of Google Text to Speech API further enhances the system's usability by converting the predicted labels into voice modules.

However, there is still room for improvement and expansion of the Signz project. For example, the current system only focuses on American sign language, but it can be expanded to recognize other sign languages such as British Sign Language or Auslan. Additionally, the system can be improved to recognize more complex gestures and facial expressions for a more holistic communication experience. Furthermore, the integration of natural language processing can enable the system to transcribe and translate the spoken language to sign language in real-time.

Overall, the Signz project provides a promising solution to improve the communication experience for Deaf and Dumb individuals. The potential for further development and expansion makes it an exciting area for future research and innovation.

References

[1] Karma Wangchuk, Panomkhawn Ri yamongkol and Rattapoom Waranusast, “Real-time Bhutanese Sign Language digits recognition system using Convolutional Neural Network”, June 2021.

[2] Ozge Mercanoglu Sincan, Hacer Yalim Keles. 2022. “Using Motion History Images With 3D Convolutional Networks in Isolated Sign Language Recognition”, February 2022.

[3] Zhibo Wang, Tengda Zhao, Jinxin Ma, Hongkai Chen, Kaixin Liu, Huajie Shao, Qian Wang, Ju Ren. 2022. “Hear Sign Language: A Real-Time End-to-End Sign Language Recognition System”, November 2020.

[4] Ahmed KASAPBAŞI, Ahmed Eltayeb AHMED ELBUSHRA, Omar AL-HARDANEE, Arif YILMAZ. “DeepASLR: A CNN based human computer interface for American Sign Language recognition for hearing-impaired individuals”, January 2022.

[5] Yao Du, Pan Xie, Mingye Wang, Xiaohui Hu, Zheng Zhao, Jiaqi Liu. “Full transformer network with masking future for word-level sign language recognition”, August 2022

[6] Sakshi Sharma, Sukhwinder Singh. “Vision-based hand gesture recognition using deep learning for the interpretation of sign language”, November 2021

[7] Boban Joksimoski, Eftim Zdravevski, Petre Lameski, Ivan Miguel Pires, Francisco José Melero, Tomás Puebla Martinez, Nuno M. Garcia, Martin Mihajlov, Ivan Chorbev, Vladimir Trajkovik. “Technological Solutions for Sign Language Recognition: A Scoping Review of Research Trends, Challenges, and Opportunities”, March 2022.

[8] Nikolas Adaloglou, Theocharis Chatzis, Ilias Papastratis, Andreas Stergioulas, Georgios Th. Papadopoulos, Vassia Zacharopoulou, George J. Xydopoulos, Klimnis Atzakas, Dimitris Papazachariou, Petros Daras. “A Comprehensive Study on Deep Learning-Based Methods for Sign Language Recognition”, April 2021

[9] Yongsen Ma, Gang Zhou, Shuangquan Wang, Hongyang Zhao, And Woosub Jung. “SignFi: Sign Language Recognition Using WiFi”, March 2018

-

[10]Bilel Ben Atitallah, Zheng Hu, Dhouha Bouchaala, Mohammed Abrar Hussain, Amir Ismail, Nabil Derbel. “Hand Sign Recognition System Based on EIT Imaging and Robust CNN Classification”, November 2022

[11] Anjan Kumar Talukdar, M. K. Bhuyan.“Vision-Based Continuous Sign Language Spotting Using Gaussian Hidden Markov Model”, June 2022

[12] Walaa Aly; Saleh Aly; Sultan Almotairi.“User-Independent American Sign Language Alphabet Recognition Based on Depth Image and PCANet Features”, September 2019

[13] Muhammad Al-Qurishi; Thariq Khalid; Riad Souissi.“Deep Learning for Sign Language Recognition: Current Techniques, Benchmarks, and Open Issues”, September 2021

[14] Tamer Shanableh, Khaled Assaleh, M.Al-Rousan.“Spatio-Temporal Feature-Extraction Techniques for Isolated Gesture Recognition in Arabic Sign Language”.May 2007

[15] Asifullah Khan, Anabia Sohail, Umme Zahoora & Aqsa Saeed Quresh.. “A survey of the recent architectures of deep convolutional neural networks”, April 2020