IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Sign Language Detection System using Machine Learning

Arunkumar N[1] , Karthik M[2], Abhinandan D L[3], Anurag Rajithran M[4], Arirajan P[5]
Assistant Professor[1], UG Student[2] ,UG Student[3] ,UG Student[4] ,UG Student[5] ,
[1,2,3,4,5] Department of Electronics and Communication Engineering, SNS College of Technology, Coimbatore, Tamil Nadu, India.

ABSTRACT:

The development of sign language detection systems leveraging machine learning techniques represents a pivotal advancement in bridging communication gaps for individuals with hearing impairments. Sign language, a visual and gestural form of communication, plays a vital role in enabling deaf and hard-of-hearing individuals to express themselves and interact with the world around them. These systems utilize computer vision and natural language processing algorithms to recognize and interpret the intricate gestures and movements that constitute sign languages. By doing so, they provide real-time translation of sign language into text or spoken language, making it accessible to a wider audience and facilitating more inclusive communication.

Key word: CNN, TensorFlow, TSL, OpenCV

INTRODUCTION:

An essential communication tool for people who are hard of hearing is sign language. The accessibility and inclusivity for this population can be greatly enhanced by creating a system that can translate sign language into text or speech. Recent advancements in sign language identification systems have been largely attributed to machine learning and computer vision technologies. This article presents the development of a Sign Language Detection System, mostly using Jupyter Notebooks within the Anaconda environment as the major coding platform, and Python as the programming language. Convolutional Neural Networks (CNNs) were constructed using key libraries such as TensorFlow and OpenCV, and data pretreatment techniques were applied. In order to recognize sign language in real time, real-time video from a webcam was also captured sing the cv2 module. By learning how to record live video from a webcam and use a CNN model to decipher sign language motions, you will make a big step toward improving communication accessibility and inclusivity for all.

RELATED WORKS:

Research efforts in the field of Sign Language Recognition (SLR) have increased significantly, and a variety of techniques and methodology have contributed to the field's accomplishments. Zhibo Wang et al. [5] introduced DeepSLR, a real-time end-to-end continuous SLR system designed to translate sign language into audible speech, addressing the challenge of making sign language accessible to a broader audience. This system leveraged sMEG and IMU sensors to capture arm movements and finger motions. R. Kothadiya et al. [9] emphasized the potential of vision-based static sign gesture recognition in human-computer interaction applications. Leveraging a multi-head attention-based encoding framework, this study achieved high accuracy with a minimal number of training layers and epochs. Furthermore, the framework's versatility extended to detecting sign gestures under varying conditions, such as different angular positions and brightness levels. Biao Xu et al. [3] presented a novel application of tensor-train factorization to Sign-to-Video Transduction (S2VT) for Chinese sign language recognition. By significantly reducing the number of model parameters, this approach enabled the deployment of SLR models on resource-constrained devices. Notably, the method-maintained accuracy in the fully-connected layer and the first LSTM layer of the S2VT model, offering a viable solution for enhancing communication for the hearing impaired.

Sunusi Bala Abdullahi et al. [8] optimized the accuracy of 3D video skeletal hand joints information for recognizing correlated American Sign Language (ASL) words. The study introduced discriminative features for Deep Bi-LSTM recognition and emphasized the need to address similar ASL words as a multi-feature problem.

Ozge Mercanoglu Sincan et al. [11] proposed novel approaches to isolated SLR, aiming to achieve robust performance with minimal computational resources. The research focused on models using RGB-MHI images and a late fusion technique for combining RGB and RGB-MHI features, achieving competitive performance with models using multiple data modalities.

EXISTING SYSTEM

There are currently very few sign language detection systems available that are especially designed for Tamil sign language. The distinct subtleties and complexities of Tamil sign language remain completely unexplored in the literature that is now available, despite the fact that several systems have been established for American Sign Language.

Due to linguistic and cultural differences, a specific method is needed, and this gap emphasizes the need for study and development in the field of Tamil sign language identification. These approaches, which attempt to decipher and translate movements within the lexicon of Tamil sign language, frequently use image recognition and deep learning models. Future efforts can greatly advance inclusive communication options for the Tamil-speaking community with hearing impairments by filling this deficiency.

PROPOSED METHODOLOGY

The proposed system for sign language detection using machine learning, specifically tailored for Tamil Sign Language, offers an efficient solution for gesture recognition and interpretation. The system begins with an input image containing a sign gesture in Tamil Sign Language. the proposed system utilizes machine learning algorithms, TensorFlow, OpenCV, Python, and Jupyter Notebook to create an end-to-end solution for Tamil Sign Language detection. This system not only identifies sign gestures in real-time images but also seamlessly translates them into corresponding textual representations, enhancing accessibility and communication for individuals using Tamil Sign Language.

The proposed research usually aims to start from image acquisition and intended to facilitate communication for the TSL community by recognizing the sign images. The algorithm we intent to use is largely based on the supervised learning approach with labeled dataset images. We choose a CNN Model to start model training the labeled dataset of images.

A. IMAGE ACQUSITION

Folder for the images are segregated separately based on train and test the dataset. Now we collect the images by capturing it from the camera to store it in the system for further processing. This provides the dataset required for the Sign Language Detection System. The collected images indicate sign gestures that depicts Tamil Sign Language. The method of show preparing starts with the procurement of pictures speaking to different sign signals in Tamil Sign Dialect. These pictures serve as the essential information input for preparing the convolutional neural organize (CNN) show. To guarantee an assorted and agent dataset, images are captured employing a webcam or other imaging gadget beneath controlled conditions, capturing diverse hand positions, lighting conditions, and foundations. The pictures are put away in assigned organizers, isolated into isolated catalogs for the preparing and testing datasets.

B. PREPROCESSING

The acquired images go through preprocessing stages to improve their quality and suitability for preparation before the model is trained. This frequently entails scaling the images to a predetermined size, cutting to align the hand location with the sign signal, and normalizing to account for variations in contrast and brightness. Additionally, data expansion techniques like turning, interpreting, and flipping may be linked to expand the dataset and strengthen the model's ability to handle variations in hand introduction and positioning.

C. LABELED DATASET

The captured images are stored in the system folder which can further be taken for labelling them to box out the actual sign gesture excluding the environment around that sign for good precision and accuracy. It is done by labelling each captured images with the letter denoting the hand sign gesture, it is done by enclosing a box around the boundaries of that sign gesture.

Following preprocessing, each sign signal is tagged with the corresponding Tamil letter or character that it corresponds to. This labeling preparation entails physically marking each image with the appropriate lesson or category, which is typically indicated by the hand gesture seen in the image. usually achieved by directly assigning a course name to the image record or by enclosing a bounding box around the hand location holding the sign motion. The annotated images are arranged into separate envelopes according to their respective groups or categories, facilitating easy access and management during the preparation process.

MODEL TRAINING

The labeled dataset images are now split in the ration of 80:20 to train and test folder respectively. This is made to ensure that the system provides good accuracy and precision with good confidence on the sign gesture that it is getting trained with. This split ensures that the show is prepared on a differing run of illustrations whereas too giving concealed information for assessing its generalization execution.

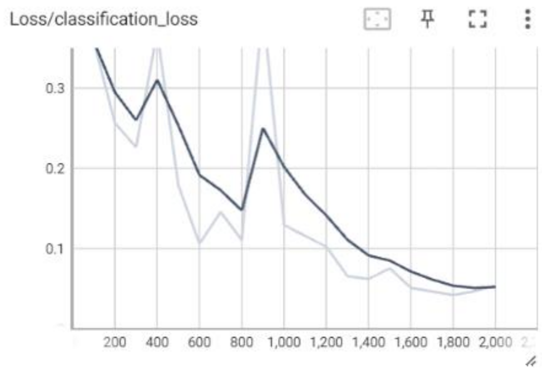
Using the images in the preparation set, the CNN show is put together during the preparation phase. The program adjusts its internal settings through an iterative optimization process as it gains the ability to identify patterns and highlights in the input images associated with each sign signal path. This preparation process entails feeding a large number of images through the organize, calculating the error between the expected yields and the actual names, and updating the weights of the model using optimization methods such as slope plunge.

The system employs TensorFlow to build and train a convolutional neural network (CNN) model. CNNs are well-suited for image recognition tasks, capturing spatial hierarchies of features. Further processing regarding model training involves the usage of Convolutional Neural Network (CNN), that requires it input images to be 320*320 in this model.

Each frame from the live webcam feed is resized to 320x320, cropped to focus on the hands, and normalized for illumination variations in real-time for immediate processing.

By using the input image size of 320*320 it provides a good balance of speed and accuracy that is need for the real time application purposes. The model generates bounding boxes and confidence scores for multiple object classes, enabling precise identification and location of various objects within an image.

This model has a box coder using faster rcnn with y scale:10.0, x scale:10.0, height scale: 5.0 and width scale:5.0. Anchor generator has 6 number of layers, minimum scale set at 0.2, maximum scale set at 0.95 and its aspect ratios are set at 1.0, 2.0, 0.5, 3.0 and 0.3333.



DATASET AND EVALUATION

The dataset of Tamil Sign Language letters consists of 31 combinations of Tamil letters consisting of 12 vowels, 1 aayutha ezhuthu and 18 consonants. Every combination refers to a distinct sign motion used in Tamil Sign Dialect, encapsulating the many social and etymological characteristics of the Tamil-speaking population. The signs are captured form a fixed positioned camera and the it consists of image captured at various environment surrounding, position and angle of hand. The trained model undergoes evaluation using the test set to assess its performance.

Metrics such as accuracy, precision, recall, and F1 score are computed to quantify the model's ability to generalize to unseen data. The basis of precision, recall, and F1-Score comes from the concepts of True Positive, True Negative, False Positive and False Negative.

```
'Loss/localization_loss': 0.016743366,
'Loss/regularization_loss': 0.14825338,
'Loss/total_loss': 0.21894392,
'learning_rate': 0.07991781}
```

ADVANTAGES:

1. Accessibility and Inclusivity: The system improves accessibility for people with hearing impairments by translating sign language motions into written representations or spoken language in real-time. This facilitates their ability to communicate more effectively with a larger audience. This encourages equality of participation and inclusiveness in a range of professional, social, and educational contexts.

2. Efficiency and Convenience: The system's real-time translation feature saves time and effort for sign language users and their communication partners by doing away with the necessity for manual interpretation or translation of sign language movements. This makes communication interactions more convenient and efficient, resulting in information transfers that are more fluid and seamless.

3. Accuracy and Reliability: The system recognizes and interprets sign language motions with high accuracy and reliability by utilizing convolutional neural networks (CNNs) and sophisticated machine learning methods. This reduces mistakes and misunderstandings by guaranteeing a constant and accurate translation of gestures into written or spoken words.

4. User-Friendly Interface: Both sign language users and communication partners will find the system's user-friendly interface to be highly intuitive and easy to use. Simple visual cues and feedback allow for smooth communication interactions even for people with little experience with assistive technology or technology in general.

RESULT:

The smooth operation of the Sign Language Detection System provides a revolutionary solution for those with hearing loss by enabling the real-time translation of Tamil Sign Language movements into spoken language or textual representations. The system transforms sign language motions into text or voice in real-time, enabling smooth user interactions, by utilizing machine learning and computer vision technology.

The system is mostly implemented with Python programming language and Jupyter Notebooks in the Anaconda environment. TensorFlow and OpenCV libraries are utilized to build Convolutional Neural Networks (CNNs) in the system. These strong algorithms add to the efficacy and dependability of the system by accurately recognizing and interpreting sign language motions.

Apart from its sturdy construction, the system has the ability to process videos in real time and records live footage from a camera by utilizing the cv2 module. This improves user accessibility to communication in real-world situations by enabling instantaneous translation of sign language gestures into written representations or spoken language.

The technology recognizes and isolates hand motions that represent Tamil Sign Language while it watches the live video feed in real time. In this procedure, hand areas are identified inside the frame, pertinent characteristics are extracted, and the features are mapped to matching spoken words or linguistic symbols.

REFERENCES:

1. Sourav Pal, Sourabh Ruhella, Sparsh Sinha, Anurag Goel, "Sign Language to Text Classification using One-Shot Learning", 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), pp.833-837, 2023.

2. Jing Zhao, Xiao Hua Li, Jennifer C. Dela Cruz, Marvin S. Verdadero, Jenette C. Centeno, Jilbert M. Novelero, "Hand Gesture Recognition Based on Deep Learning", 2023 International Conference on Digital Applications, Transformation & Economy (ICDATE), pp.250-254, 2023.

3. Biao Xu; Shiliang Huang; Zhongfu Ye Application of Tensor Train Decomposition in S2VT Model for

Sign Language Recognition IEEE Access (Volume: 9) 2021.

4. P. Rawat, L. Kane, M. Goswami, A. Jindal, and S. Sehgal, "A Review on Vision-based Hand Gesture Recognition Targeting RGB-Depth Sensors", International Journal of Information Technology & Decision Making (IJITDM), vol. 22, no. 01, pp. 115-156, 2023.

5. Zhibo Wang; Tengda Zhao; Jinxin Ma; Hongkai Chen; Kaixin Liu; Huajie Shao; Qian Wang; Ju Ren Hear Sign Language: A Real-Time End-to-End Sign Language Recognition System, IEEE Transactions on Mobile Computing (Volume: 21, Issue: 7, 01 July 2022)

6. Wadhawan, A., Kumar, P. Deep learning-based sign language recognition system for static signs. Neural Comput & Applic Volume 32, 7957–7968 (2020).

7. S.A.M.A.S Senanayaka; R.A.D.B.S Perera; W. Rankothge; S.S. Usgalhewa; H.D Hettihewa; P.K.W. Abeygunawardhana Continuous American Sign Language Recognition Using Computer Vision And Deep Learning Technologies, 2022 IEEE Region 10 Symposium (TENSYMP)

8. Sunusi Bala Abdullahi; Kosin Chamnongthai American Sign Language Words Recognition Using Spatio-Temporal Prosodic and Angle Features: A Sequential Learning Approach IEEE Access (Volume: 10), 2022.

9. R. Kothadiya; Chintan M. Bhatt; Tanzila Saba; Amjad Rehman; Saeed Ali Bahaj SIGNFORMER: DeepVision Transformer for Sign Language IEEE Access (Volume: 11),2023.

10. Nikolas Adaloglou; Theocharis Chatzis; Ilias Papastratis; Andreas Stergioulas; Georgios Th. Papadopoulos A Comprehensive Study on Deep Learning-Based Methods for Sign Language Recognition and IEEE Transactions on Multimedia (Volume: 24) 2021.

11. Ozge Mercanoglu Sincan; Hacer Yalim Keles Using Motion History Images With 3D Convolutional Networks in Isolated Sign Language Recognition IEEE Access (Volume: 10) 2022.

12. Jungpil Shin; Abu Saleh Musa Miah; Kota Suzuki; Koki Hirooka; Md. Al Mehedi Hasan Dynamic Korean Sign Language Recognition Using Pose Estimation Based and Attention-Based Neural Network IEEE Access (Volume: 11),2023.

13. Vasileios Kouvakis; Stylianos E. Trevlakis; Alexandros-Apostolos A. Boulogeorgos Semantic Communications for Image-Based Sign Language Transmission IEEE Open Journal of the Communications Society (Volume: 5), 2024.

14. David Richard Tom Hax; Pascal Penava; Samira Krodel; Liliya Razova; Ricardo Buettner A Novel Hybrid Deep Learning Architecture for Dynamic Hand Gesture Recognition IEEE Access (Volume: 12), 2024.

15. Abu Saleh Musa Miah; Md. Al Mehedi Hasan; Satoshi Nishimura; Jungpil Shin Sign Language Recognition Using Graph and General Deep Neural Network Based on Large Scale Dataset IEEE Access (Volume: 12),2024.

16. Zhenxing Zhou; Vincent W. L. Tam; Edmund Y. Lam SignBERT: A BERT-Based Deep Learning Framework for Continuous Sign Language IEEE Access (Volume: 9),2021.