Available online at http://www.journalijdr.com

IJDR

ISSN: 2230-9926

International Journal of Development Research
Vol. 13, Issue, 03, pp. 62185-62188, March, 2023
https://doi.org/10.37118/ijdr.26518.03.2023

RESEARCH ARTICLE OPEN ACCESS

A NOVEL APPROACH TO RECOGNISING SIGN LANGUAGE USING DEEP LEARNING TECHNIQUES

B. S. Panda[1]*, R Kalpana[2], P Hema Latha[2] and S Sofia[2]

[1]Professor, Department of CSE, Raghu Engineering College, Visakhapatnam, Andhra Pradesh, India
[2]Students, Department of CSE, Raghu Engineering College, Visakhapatnam, Andhra Pradesh, India

ARTICLE INFO

Article History:
Received 18th January, 2023
Received in revised form
09th February, 2023
Accepted 21st February, 2023
Published online 30th March, 2023

KeyWords:
Long-Short Term Memory (LSTM), TensorFlow, Indian Sign Language (ISL), MediaPipe, Keras, Adam optimizer.

Corresponding author: B. S. Panda

ABSTRACT

A new technology has been developed to assist communication between sign language users and non-sign language speakers, specifically for Indian Sign Language (ISL). This technology uses computer vision and deep learning techniques to translate ISL into text. The system captures signs using a camera and maps each sign to its meaning through trained data before converting it into text using TensorFlow. OpenCV and MediaPipe are used to extract keypoints from real-time frames for specific sign language actions, which are then preprocessed by normalizing to a mean of zero, extracting relevant features using MediaPipe, and encoding labels with one-hot encoding. Keypoints from different body parts are concatenated into a single feature vector containing position and visibility information. An LSTM model is created and trained on the preprocessed data using TensorFlow, with the Adam optimizer and categorical cross-entropy loss. The model's hyperparameters are tuned, and regularizes are used to enhance its performance. This system has demonstrated high accuracy in recognizing ISL and converting it into text, which has the potential to be used in various industries, including healthcare, gaming, and education.

Citation: ***B. S. Panda, R Kalpana, P Hema Latha and S Sofia. 2023.*** "A novel approach to recognising sign language using deep learning techniques". *International Journal of Development Research*, 13, (03), 62185-62188.

INTRODUCTION

Sign language is a form of communication that does not rely on speech and is employed by individuals who are deaf or mute. It is globally recognized and allows people to express themselves without the need of verbal menas. There are various sign languages, including Indian Sign Language(ISL) [1], American Sign Language(ASL)[2,3], and Chinese Sign Languag(CSL)[4,5]. A technology using computer vision and deep learning can translate Indian Sign Language (ISL) into text by mapping signs to meanings and using TensorFlow to convert them.This technology has shown impressive accuracy in recognizing ISL [6] and converting it to text, providing an encouraging solution to improve communication access for the hard and deaf of hearing community in India. It has potential applications in healthcare, education, gaming, and various other industries.The development of this technology using computer vision and deep learning offers a promising solution to improve communication accessibility for the hard and deaf of hearing community in India. By translating Indian Sign Language into text with high accuracy, this technology can help bridge the communication difference between sign language users and non-sign language speakers. After introducing the topic of the research paper, the literature survey gathers information on the research topic, and the methodology section explains the research design, data collection methods, and analytical techniques. The algorithm section outlines the steps taken to develop a specific algorithm, while the result analysis presents the findings, often using tables and graphs. The discussions section analyzes the results, identifies limitations, and suggests areas for future research, and the conclusion summarizes the findings, their relevance, and highlights contributions and limitations.

LITERATURE SURVEY

Saurabh Kumbhar etal.,[7] suggested method for recognizing sign language alphabets and numbers, which utilizes Convolutional Neural Networks (CNN) .This system has been developed using recent deep learning technology. CNN is used due to its ability to recognize hidden patterns and correlations in raw data, leading to increased system accuracy. This system is designed to assist individuals with hearing and language impairments. However, the CNN model is more complicated for constructing and deploying. Input images are processed by the system to identify and interpret sign language. Yogeshwar I. Rokade etal., [8] suggested a technique to identify the symbols of ISL using automatic recognizer, discussed in this paper, which has potential to benefit individuals with hearing and speaking disabilities. The system involves several steps such as skin color segmentation, binary image transformation, Euclidean distance transformation, row and column projection, and feature extraction with central and HU moments. Classification is done using neural networks and SVM. Kshitij Bantupalli etal.,[9] focus of the proposed work is to develop a vision-based application that facilitates communication among individuals proficient in sign language and those who do not understand sign language. The system employs deep learning and computer vision techniques to recognize and translate sign language gestures into text. The proposed model is capable of processing video sequences and extracting features from them. Inception, a Convolutional Neural Network (CNN), is developed for recognizing spatial features, while a Recurrent Neural Network (RNN) is trained to identify adhoc features. To train the model, the American Sign Language Dataset is utilized. Umang Patel etal.,[10] presents a solution to the communication barrier between normal people and those who are deaf and dumb, by converting sign language to text and speech. Hand gestures are captured and processed using MATLAB, and features are evaluated using a technique called moment. PNN and KNN techniques are used to convert the gestures into English and Hindi. The paper concludes that this method helps to facilitate communication between normal and hearing-impaired individuals. Suharjito etal.,[11] proposed the complexity and variation of Sign Language make it difficult to recognize accurately. Researchers have tried various methods to improve recognition, including using Action Recognition models like i3d inception. The main objective of this paper is to employ the i3d inception model to transfer learning for Sign Language Recognition. The model achieved 100% accuracy on training but had low validation accuracy and was overfit. The studies mentioned above suggest that sign language recognition technology has the potential to improve accessibility and inclusivity for individuals who are hard or deaf of hearing [12]

METHODOLOGY

The development of sign language recognition systems involves data collection, preprocessing, feature extraction, and model training. Machine learning and computer vision have enabled the creation of systems that can convert Indian Sign Language (ISL) into spoken or written language. We propose a deep learning-based approach for developing an ISL recognition system. In this methodology includes dataset collection, preprocessing, feature extraction, and LSTM [14] network design and training. We also evaluate the model performance using standard metrics.

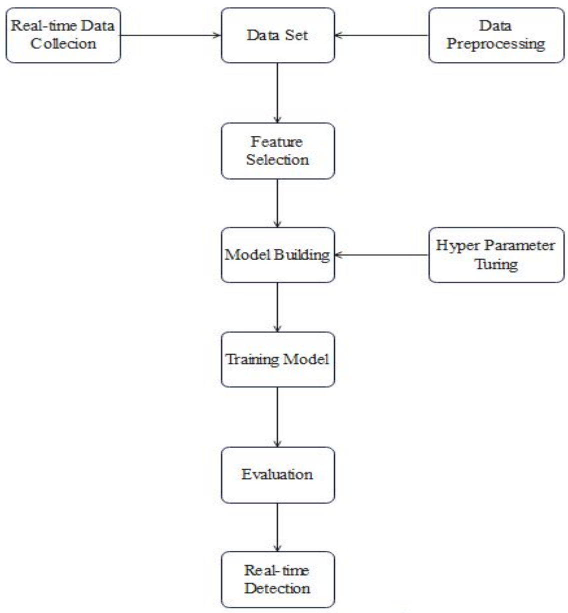

Fig. 1. Methodology

Dataset: The dataset used in this project is Indian Sign Language alphabets. Keypoints were extracted using OpenCV and Mediapipe from frames captured in real-time. Directories were created for each sign and the keypoints were saved as numpy files for each frame. It contains signs and some words. This created a dataset for training and testing sign language recognition models.

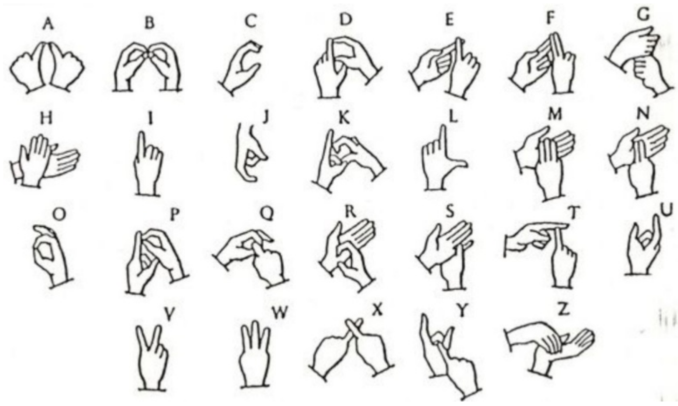

Fig. 2. Data used for training

A. ***Data collection:*** We employed OpenCV and Mediapipe to detect and extract keypoints from real-time frames capturing specific actions (signs) to gather data. We created directories for each action and saved the keypoints (landmarks of face, pose, left hand and right hand) as numpy files in the corresponding directory for each frame.
B. ***Preprocessing:*** We preprocessed the data by normalizing it to a mean of zero, extracting relevant features using Mediapipe library, and encoding labels using one-hot encoding.
C. ***Feature extraction:*** Keypoints are extracted from different body parts including pose, face, left hand and right hand using Mediapipe and concatenated into a single feature vector containing position and visibility information.
D. ***Model creating and training:*** An LSTM model with 3 layers is trained on X_train and y_train, divides into training and validation sets. This sets consisting, test size of 0.32 and a random state of 42. This model has 64, 128, and 64 units in the three LSTM layers, respectively, followed by three dense layers. This model uses the Adam optimizer with a learning rate of 0.0001 and categorical cross-entropy loss. Categorical accuracy is used to evaluate the model, and TensorBoard is used to log the training progress.

Algorithm

The algorithm used in this is LSTM MODEL [15] stands for Long Short Term Memory, which is a method of RNN (Recurrent Neural Network). The purpose of its design is to overcome the problem of vanishing gradient encountered in conventional RNNs, which limits its capacity to learn dependencies in consequent data that occur over extended periods. LSTM networks [16. 17] use a specialized memory cell and uses three gates - input, forget, and output, for controlling the flow of information through the network selectively. This enables them to learn and remember dependencies in consequent data that occur over extended periods, particularly suitable for tasks like recognition of speech, processing of natural languages, and prediction of time series.

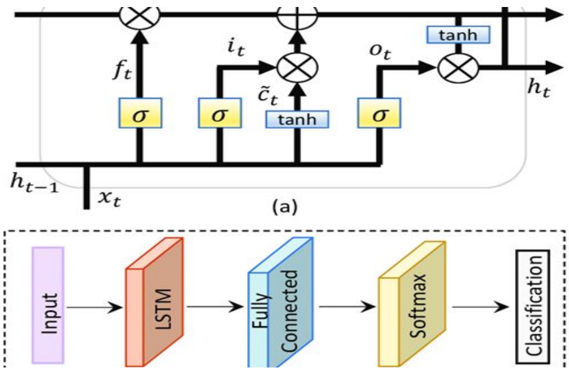

Fig. 3. LSTM Model

RESULT ANALYSIS

Table 1. Result Analysis-1

Algorithm	Accuracy
ANN	0.9437
LSTM	0.9828

The variation in accuracy indicates that the LSTM model was better equipped to identify intricate patterns and interdependencies in the data as compared to the ANN model. This could be due to the fact that LSTMs belong to a category of recurrent neural networks that are specifically tailored for processing sequential data, whereas ANNs are better suited for simpler, non-sequential data.

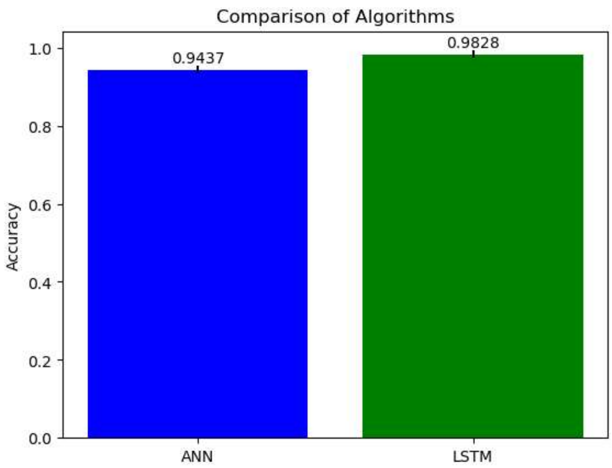

Fig. 4. Comparison of algorithm

The model got Precision as 0.9828 and F1 score as 0.9828.

Table 2. Result Analysis-2

Evaluation Metric	Model Performance
Accuracy	0.9828
Precision	0.9828
F1-score	0.9828

The precision and F1 score of the model are the same (i.e., 0.9828), it suggests that the model has performed well on the positive class. Precision is the fraction of true positive predictions out of all positive predictions, while F1 score is a weighted average of precision and recall. Therefore, when precision and F1 score are the same, it means that the model has correctly predicted a high percentage of true positives while maintaining a low false positive rate. Overall, these performance metrics indicate that this model is making accurate predictions and is performing well on the given dataset.

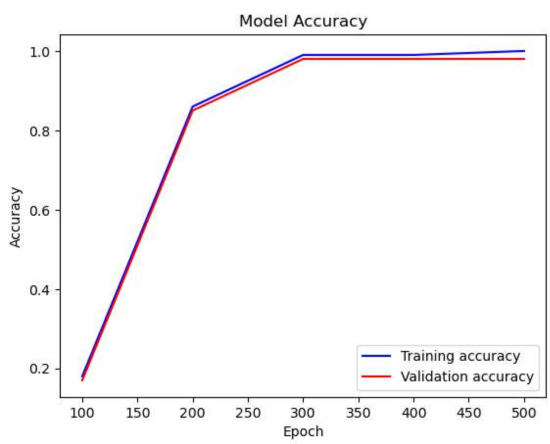

Fig. 5. Training and Validation accuracy

DISCUSSIONS

During the process of training and evaluating the models, we made several observations. Firstly, we noticed that increasing the values of certain parameters in MediaPipe resulted in different outputs, indicating that the results of the models can be affected by the input data. We also experimented with the number of epochs used to train the models, and found that training the models for more epochs resulted in better accuracy. However, we also needed to balance this with the risk of overfitting the model to the training data. To ensure consistency in our results, we set the random seed for the data splitting process so that the same data was implemented for training and testing each time the model was run. To improve the stability and performance of the models, we used various techniques such as dropout, L2 regularization, and batch normalization. We also set the learning rate of the Adam optimizer to 0.0001, which helped the models converge more accurately. Overall, these observations and techniques were used to help us achieve the highest possible accuracy while ensuring that the models were stable and not overfitting to the training data. The limitations include overfitting, dependence on the dataset, and hyperparameter tuning.

CONCLUSION

This study aimed to develop a deep learning model using LSTM for recognizing Indian sign language. According to the results LSTM model outperformed than the ANN model and other machine learning algorithms, indicating its suitability for handling sequential data with complex patterns and dependencies. The high precision and F1 scores suggest that the LSTM model is making accurate predictions while capturing most of the positive cases in the dataset. However, the model has some limitations, including overfitting, dependence on the dataset, and hyperparameter tuning. To improve its performance, fine-tuning the hyperparameters and using techniques such as cross-validation may be necessary. Overall, the LSTM model shows promise for multi-label classification tasks on sequential data. Future research could explore the design of a model to detect motion signs, which could be an interesting area for further investigation.

REFERENCES

American Sign Language Recognition using Deep Learning and Computer Vision by Kshitij Bantupalli and Ying Xie, 2018 IEEE International Conference on Big Data (Big Data).

C. Charayaphan and A. Marble, “Image process1ing system for interpreting motion in american sign language,” Journal of Biomedical Engineering, vol. 14, no. 5, pp. 419–425, 1992.

Deaf Cambridge Dictionary. (2018). Retrieved from Cambridge Dictionary: https://dictionary.cambridge.org/dictionary/english/deaf

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A Large-Scale Hierarchical Image Database. Computer Vision and Pattern Recognition, CVPR 2009. IEEE Conference on (pp. 248-255). IEEE. Miami, FL, USA.

F. A. Gers, J. Schmidhuber, and F. Cummins, (2000) “Learning to forget: Continual prediction with LSTM,” Neural Computation, vol. 12, no. 10, pp. 2451–2471.

F. A. Gers, N. N. Schraudolph, and J. Schmidhuber, (2003) “Learning precise timing with LSTM recurrent networks,” Journal of Machine Learning Research, vol. 3, pp. 115–143.

Indian Sign Language Recognition System by Saurabh Kumbhar, Abhishek Landge, Akash Kulkarni, Devesh Solanki, Vidya Kurtadikar, International Journal of Innovative Science and Research Technology ISSN No:-2456-2165

Indian Sign Language Recognition System by Yogeshwar I. Rokade, Prashant M. Jadav, DOI: 10.21817/ijet/2017/v9i3/170903S030.

J. Ma, W. Gao, J. Wu, and C. Wang, “A continuous chinese sign language recognition system,” in Au1tomatic Face and Gesture

Recognition, 2000. Pro1ceedings. Fourth IEEE International Conference on. IEEE, 2000, pp. 428–433.

Learned Sign Language from: https://indiansignlanguage.org/

Moment Based Sign Language Recognition for Indian Languages by Umang Patel and Aarti G. Ambekar,2017 Third International Conference on Computing, Communication, Control And Automation (ICCUBEA).

S. Lang, M. Block, and R. Rojas, "Sign language recognition using kinect," in Artificial intelligence and soft computing. Springer, 2012, pp. 394–402.

Sign Language Recognition Using Modified Convolutional Neural Network Model by Suharjito, Herman Gunawan, Narada Thiracitta and Ariadi Nugroho, The 1st 2018 INAPR International Conference, 7 Sept 2018, Jakarta, Indonesia.

T. Starner and A. Pentland, "Real-time american sign language recognition from video using hid1den markov models," in Motion-Based Recognition. Springer, 1997, pp. 227–243.

T. Takahashi and F. Kishino, "Hand gesture coding based on experiments using a hand gesture interface device," SIGCHI Bull., vol. 23, no. 2, pp. 67–74, Mar. 1991. [Online]. Available: http://doi.acm.org/10.1145/122488.122499.

W. Shenghui, C. Weihao, and C. Kuanglin, (2015) "Design and implementation of sign language interpreter system based on leap motion (in chinese)," Modern Electronics Technique, vol. 38, no. 24, pp. 114– 117.

Weichert, D. Bachmann, B. Rudak, and D. Fis1seler, (2013)"Analysis of the accuracy and robustness of the leap motion controller," Sensors, vol. 13, no. 5, pp. 6380–6393.

****