Technical Answers for Real World Problems

CSE1901

Topic Name: Detection of Indian Sign Language using Convolutional Neural Network with Attention layer

Domain: Deep Learning

Team Details:

Aditya Kalra 19BCE0390
Parth Mahadik 19BCE0443
Sarthak Bharadwaj 19BCE0859
Ayush Bansal 19BCE0862

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Abstract:

Communication is one of the major areas where differently abled (deaf and dumb) people face problems on everyday basis. Recent advances in technology have helped in development of Sign Language Recognition systems and models which has helped in solving this problem to a great extent. Machine Learning based Sign Language Recognition models have shown to have high accuracy. However due to the large volume of data required to train the model, the training can be considerably long. To counter this problem, we propose an accelerated Indian Sign Language Recognition Model using Distributed Learning. The model uses Convolutional Neural Network (CNN) for recognizing signs and converting them to English language. Graphical Processing Units (GPUs) have been utilized to make the model training process more efficient.

Literature Survey:

S No.	Paper Title	Methodology	Needed Improvement	Future Work

1.	Real Time Sign Language Recognition Using Image Classification, H. Lakhotiya, H. S. Pandita and R. Shankarmani	The model proposed in this paper is implemented for the American sign language, it also proposes a real time image detection and processing application, which proves that real-time hand signal detection and classification is a viable endeavour.	The algorithm fails with a noisy background which can be improved.	Current implementation only processes input from plain background, and the authors declare in their future scope that they aim to augment the pre-processing algorithm to detect gestures in noisy backgrounds as well.

2.	Convolutional Neural Network (CNN) for Image Classification of Indonesia Sign Language Using TensorFlow.	This paper proposes a CNN model to detect and identify various hand signals part of the Indonesian hand sign language. It elaborates upon the process of development of the model, along with robust experimental results. The proposed model achieved an accuracy of 96.67% on the training dataset, proving that such a classification model is indeed viable	The accuracy of the model can be improved and the range of characters recognised by the algorithm can also be increased.	We can improve the model and increase the range of characters recognised by it, we can improve on the gesture detection and identification aspect .

		and can be implemented in the context of Indian hand sign language .		
3	Achieving Real-Time Sign Language Translation Using a Smartphone's True Depth Images	This paper proposes a system that can detect and classify sign language symbols in real-time from smartphone processed images, thus providing us with the confidence that today's smartphones have sufficient computational power in order to translate sign language to plain text and back in real-time from	The algorithms can be improved to work on different situations like change in lighting etc efficiently.	Since most of the experiments in the proposed work was only done for one type of depth sensor, the authors plan to work on RGB-based depth sensors, and to analyse the changes in performance in different situations such as change in lighting, etc.

		smartphone quality images.		

4	Towards Multilingual Sign Language Recognition.	This paper focuses on multilingual sign language. The hand movement modelling was done with usage of target sign language independent dataset by derivation of subunits of hand movement. The proposed approach was validated against different types of Sign Language. It demonstrates that sign language recognition models could be developed by	We can address resource constraint issues in sign language processing such as, developing systems with reduced number of signers and examples to improve the algorithm.	The future work will build upon these finding to address resource-constraint issues in sign language processing such as, developing systems with reduced number of signers and examples.

		utilizing multilingual sign language data. Although considerable performance difference has been observed when hand modelling is done in a language independent manner rather than in language dependent manner.		

5	SRHandNet: Real-Time 2D Hand Pose Estimation With Simultaneous Region Localization	This paper proposes a novel method of optimization of hand gestures by iteratively increasing the number of Regions of Interest (ROIs) calculated by the model in order to boost the accuracy as and when needed. This provides yet another way to detect and classify hand gestures in real-time with a relatively low computational cost.	The accuracy of the algorithm can be improved.	The researchers aim to improve their hand gestures recognition algorithm and weed out its limitations

6	Indian Sign Language Gesture Recognition using Image Processing and Deep Learning	This paper proposes a realtime model for ISL gesture recognition, based on the incoming image data from the Kinect. Effective real time background subtraction was done using depth perception techniques. Computer vision techniques were used to achieve one-to-one mapping between the depth and the RGB pixels		For future work the research can be focused on real-time prediction of more words related to ISL and also on sentence formation.

7	Benchmarking deep neural network approaches for Indian Sign Language recognition	In this paper, an extensive comparative analysis of various gesture recognition techniques involving convolutional neural networks and machine learning algorithms has been discussed and tested for real-time accuracy. Three models: a pre-trained VGG16 with finetuning, VGG16 with transfer learning and a hierarchical neural network were analysed based on a number of trainable parameters. These models	The models mentioned in only deal with images and need to be modified to capture frames.	For future work a system that is capable of generating sentences for better practical usages can be developed and more emphasis can be given on optimization of current algorithms to further improve the model to learn the essential features and backgrounds.

		were trained on a self-developed dataset consisting images of Indian Sign Language (ISL) representation of all 26 English alphabets		

8	Motionlets Matching With Adaptive Kernels for 3-D Indian Sign Language Recognition	This paper proposes characterizatio n of sign language gestures articulated at different body parts as 3D motionlets, which describe the signs with a subset of joint motions. The proposed method is sign invariant to temporal misalignment and can characterize sign language based on a 3D spatio-tempora l framework.	Three feature kernels based on trajectories, finger shape and their orientations are constructed, which measure the similarity between the query signs and the database signs	No future Scope was suggested by the authors other than increasing its accuracy and precision compared to the proposed work.

9.	American Sign Language Recognition Using Leap Motion Controller with Machine Learning Approach	This paper presents an American Sign Language recognition system which involves 26 letters and 10 digits using the Leap Motion Controller. This paper indicates that the distance between one fingertip and the adjacent fingertips is a significant feature for sign language recognition.	Lower mean accuracy rate due to the large similarity between certain letters and digits.	Future works should also consider expanding the sign language recognition system to word- and sentence-based recognition, as well as to other languages, instead of limiting recognition to only ASL

10	Recent advances in deep learning for object detection.	By reviewing a large body of recent related work in literature, this paper systematically analyses the existing object detection frameworks and organize the survey into three major parts: (i) detection components, (ii) learning strategies, and (iii) applications & benchmarks.		Scalable, Proposal Generation and Encoding of Contextual Information

11	Deep learning-based sign language recognition system for static	This paper deals with robust modelling of static signs in the context of sign language recognition using deep learning-based convolutional neural networks (CNN).The efficiency of the proposed system is evaluated on approximately 50 CNN models. The results are also evaluated on the basis of different optimizers, and it has been observed that the proposed approach has achieved the highest training		For future work, there is a need to collect more datasets to refine the recognition method. In addition, the system will be extended to recognize dynamic signs which require the collection and development of a videobased dataset and the system is tested using CNN architecture by dividing the videos into frames

		accuracy of 99.72% and 99.90% on coloured and grayscale images, respectively.		
12	Static Hand Gesture Recognition using Convolutional Neural Network with Data Augmentation.	This paper has used CNN model to classify the images as recent research has proved the supremacy of Convolutional Neural Network (CNN) for image representation and classification. Since, CNN can learn complex and nonlinear relationships among images, in this paper, a static hand gesture recognition	The project emphasizes to recognize only static gestures.	It was assumed that the background should be less complex. Therefore, recognition of gestures in complex background can be another future work. Recognition of gestures made with both hands is not possible by the system. Therefore, another future work can be the recognition of gestures made with both hands.

		method deploying CNN was proposed. It also analyses the effect of data augmentation in deep learning.		
13	Development of an Infrared-Based Sensor for Finger Movement Detection	Movement Detection using infrared cameras, hand movement s detected using Haar like features and CNN		Development of a suitable hybrid model for CNN type accuracy and Haar feature like speed for optimal results.

14	Object detection with deep learning: A review	The review begins with a brief introduction on the history of deep learning and its representative tool, namely, the convolutional neural network. Then, it focuses on typical generic object detection architectures along with some modifications and useful tricks to improve detection performance further.		Different techniques must be implemented depending on the need, i.e. classification, feature extraction etc, and application, such as face detection, generic object detection,

15	Application of deep learning for object detection.	This paper demystifies the role of deep learning techniques based on convolutional neural network for object detection. Deep learning frameworks and services available for object detection are also enunciated.		Developing mechanisms to provide Object Detection as a Service

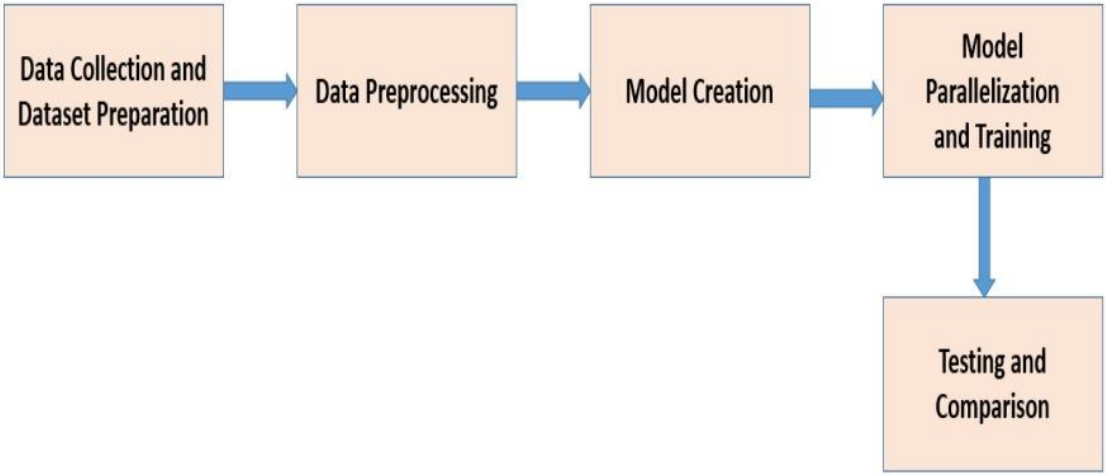
16	Sign Language Recognition Using Convolutional Neural Networks	Convolutional neural networks can be used to accurately recognize different signs of sign language, with users and surroundings not included in the training set. This generalization capacity of CNNs in spatiotemporal data can contribute to the broader research field on automatic sign language recognition. CNNs are inspired by the visual cortex of the human brain.		Capacity of CNNs in spatiotemporal data can contribute to the broader research field on automatic sign language recognition.

17	Static Sign Language Recognition Using Deep Learning	Translate static sign language into its corresponding word equivalent that includes letters, numbers, and basic static signs to familiarize the users with the fundamentals of sign language. This system is well-matched with the existing systems, given that it can perform recognition at the given accuracy with larger vocabularies and without an aid such as gloves or hand markings.		Future works should focus on improving accuracy

18	Comparing ANN, SVM, and HMM based Machine Learning Methods for American Sign Language Recognition using Wearable Motion Sensors	ANN-based machine learning methods outperforms the SVM and the HMM methods based on the overall recognition accuracy results. ANN method gives the overall best accuracy in recognizing the ASL words, but HMM method can be used to support the recognition of complete sentences in ASL.		The research presented in this paper can further be expanded by increasing the dictionary size of gestures while also incorporating non-manual markers such as head tilts, facial gestures as well as shoulder movements in the gesture recognition framework. Furthermore, future research in this field can combine existing techniques with vision-based systems to improve the classification of the data. Future research directions include

				continuous ASL recognition of full sentences using HMM.

Proposed System Architecture:



Modules:

1. Data Collection and Preparation
2. Data Preprocessing
3. Model Creation
   3.1 Model Parallelization using GPU
   3.2 Model Parallelization using TPU
4. Testing and Comparison

Programming Languages/ Tools Used:

1. Python
2. OpenCV
3. TensorFlow
4. Keras
5. Sklear
6. Numpy

Implementation Results:

Collection and Preparation of Dataset:

In the dataset we have 1200 images of each letter divided into the test and training portion.The size of the data set is enough for proper training and testing of the model that we are trying to create. After this we have compressed our data set into a zip file for use in the Colab notebook.