Real-Time Sign Language Detection and Recognition

A Major Project Synopsis Submitted to

Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal
Towards Partial Fulfillment for the Award of

Bachelor of Technology
(Computer Science and Engineering)

Guided by:

Prof.Narendra Pal Singh Rathore
Computer Science & Engineering Department
AITR, Indore

Submitted by:

Sarthak Sharma (0827CS181187)
Preet Kaur Nagi (0827CS181154)
Rahul Ahuja (0827CS181160)
Poorti Rajani (0827CS181144)

Department of Computer Science and Engineering
Acropolis Institute of Technology & Research, Indore
JAN-JUN 2022

Abstract

Sign Language detection by technology is an overlooked concept despite there being a large social group which could benefit by it. There are not many technologies which helpin connecting this social group to the rest of the world. Understanding sign language is one of the primary enablers in helping users of sign language communicate with the restof the society. So in this project we will be creating a Sign Language Detection System using OpenCV, Machine Learning and Deep Learning which will detect Sign Language gestures in Real-time, which will be helpful for deaf and disabled people.

Introduction of the Project

There have been several advancements in technology and a lot of research has been done to help the people who are deaf and dumb. Aiding the cause, Deep learning, and computer vision can be used too to make an impact on this cause. This can be very helpful for the deaf and dumb people in communicating with others as knowing sign language is not something that is common to all, moreover, this can be extended to creating automatic editors, where the person can easily write by just their hand gestures. In this sign language recognition project, we create a sign detector, which detects alphabets from A to Z that can very easily be extended to cover a vast multitude of various types of signs and hand gestures.

The sign language is used widely by people who are deaf-dumb these are used as a medium for communication. Which is nothing but an assortment of various gestures formed by different shapes of hand, its movements, orientations as well as the hand gestures. `Deaf' people have very little or no hearing ability. They use sign language for communication. People use different sign languages indifferent parts of the world.
Communication is one of the basic requirements for survival in society. Deaf and dumb people communicate among themselves using sign language but normal people find it difficult to understand their language. Extensive work has been done on American sign language recognition but Indian sign language differs significantly from American sign.

Objective

The objective of the Sign Language recognition project is to successfully create a sign detector ,which will detect alphabets from A to Z that can very easily be extended to cover a vast multitude of other signs and hand gestures including numbers also. Basically the aim of the project is to build a machine learning model that can accurately and efficiently be able to classify various hand gestures used for fingerspelling in sign language. And also learn how to use OpenCV for Real-time detection and Tensorflow/Keras to build a Model with high accuracy.

The project aims to aid individuals with inherent or acquired hearing disability facilitate real-time communication with deaf with high accuracy. We will develop this project using OpenCV and Keras modules of python. Main Objective of this Project is to create a high

accuracy classification model which should give accurate predictions in training data as well as in testing data.

Scope

Since deaf people are usually deprived of normal communication with other people, they have to rely on an interpreter or some visual communication. Now the interpreter can not always be available, so this project can help eliminate the dependency on the interpreter.
The system can be extended to incorporate the knowledge of standard sign language.
A mobile and web based version of the application will increase the reach to more people.
Integrating hand gesture recognition system using computer vision for establishing 2- way communication system.

Study of Existing System

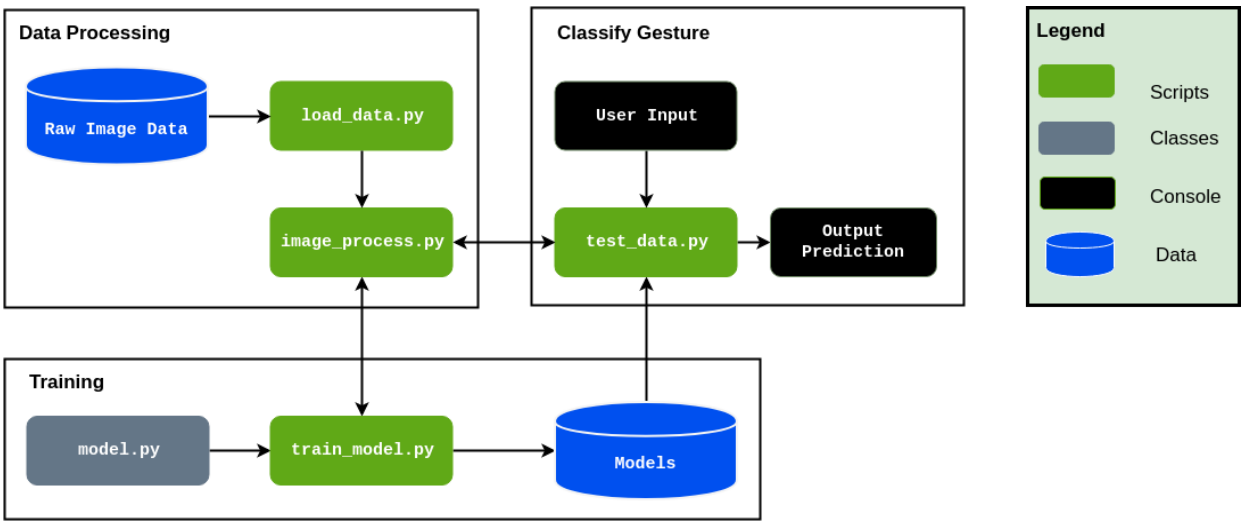
We have gone through other similar works that are implemented in the domain of sign language recognition.The summaries of each of the project works are mentioned below:

- A Survey of Hand Gesture Recognition Methods in Sign Language Recognition.
- Communication between Deaf-Dumb People and Normal People
- Intelligent Sign Language Recognition Using Image Processing
- Sign Language Interpreter using Image Processing and Machine Learning.

Project Description

On the collected dataset, we divided our approach to tackle the classification problem into three stages. The first stage is to segment the skin part from the image, as the remaining part can be regarded as noise w.r.t the character classification problem.
The second stage is to extract relevant features from the skin segmented images whichcan prove significant for the next stage i.e learning and classification.
The third stage as mentioned above is to use the extracted features as input into various supervised learning models for training and then finally use the trained models for classification.

Methodology/Planning of the Project work



Let's understand the flow of our project: **Data Processing, Training,Classify Gesture.**

Data Processing: The load data.py script contains functions to load the Raw Image Data and save the image data as numpy arrays into storage. The process data.py script will load the image data from data.py and preprocess the image by resizing the image, and applying filters and ZCA whitening to enhance features. During training the processed image data was split into training, validation, and testing data and written to storage. Training also involves a load dataset.py script that loads the relevant data split into a Dataset class. For use of the trained model in classifying gestures, an individual image is loaded and processed from the system.

Training: The training loop for the model is contained in train model.py. The model is trained with hyperparameters obtained from a config that lists the learning rate, batch size, image filtering, and number of epochs. The configuration used to train the model is saved along with the model architecture for future evaluation and tweaking for improved results. Within the training loop, the training and validation datasets are loaded as Dataloaders and the model is trained using Adam Optimizer with Cross Entropy Loss.
The model is evaluated every epoch on the validation set and the model with best validation accuracy is saved to storage for further evaluation and use. Upon finishing training, the training and validation error and loss is saved to the disk, along with a plot of error and loss over training.

Classify Gesture: After a model has been trained, it can be used to classify a new ASL gesture that is available as a file on the system. The user inputs the file path of the gesture image and the test data.py script will pass the file path to process data.py to load and preprocess the file the same way as the model has been trained.

Expected Outcome

- To provide information access and services to deaf people in American sign language.
- To develop a scalable project which can be extended to capture the whole vocabulary of ASL through manual and non manual signs.

The results on this project should have at least 95% accuracy rate. This means, if the user who is supposed to use this project has already contributed to our dataset earlier, the system will guarantee about 95% recognition rate.

Resources and Limitations :

- Camera
- Capable Processor
- Python (3.7.4)
- IDE (Jupyter)
- Numpy (version 1.16.5)
- cv2 (openCV) (version 3.4.2)
- Keras (version 2.3.1)
- Tensorflow (as keras uses tensorflow in backend and for image preprocessing) (version 2.0.0)

Models described in the given literature give poor results if the dataset includes faces of signers as the model ends up training incorrect features; the same problem occurs with the color of the background. These models also face problems if they are trained on color images and the skin tone in testing images differs from training images. While working with videos, the models take a lot of time to predict sign and the dumb people are habituated with sign language so their speed cannot be matched with these existing systems.

The Indian Sign Language lags behind its American Counterpart as the research in this field is hampered by the lack of standard datasets. Unlike American Sign Language, it uses both hands for making gestures which leads to occlusion of features. ISL is also subject to variance in locality and the existence of multiple signs for the same character. Also some characters share the same alphabet (E.g V and 2 have the same sign, similarly W and 3 have the same sign) and the resolution of the sign is context dependent.

Conclusion

In this Project, we have gone through an automatic sign language gesture recognition system in real-time, using different tools. Although our proposed work expected to recognize the sign language and convert it into the text which can be useful for deaf and dumb people there’s still a lot of scope for possible future work.

In this project, a survey on sign language recognition is presented and various techniques have been studied and analysed for the same. In the recognition process, segmentation plays a crucial part in which skin region is separated from the background which usually affects the recognition accuracy. Besides segmentation, classification also depends on the feature extraction techniques which performs dimensionality reduction and reduces the computation cost. Study of various classification techniques concludes that deep neural networks (CNN, Inception model, LSTM) performs better than traditional classifiers such as KNN and SVM.

References

[1] Refers to https://data-flair.training/blogs/sign-language-recognition-python-ml-opencv/

[2] D. Maturana and S. Scherer, "VoxNet: A 3D Convolutional Neural Network for Real-Time Object Recognition," in *IEEE International Conference*, 2015.

[3] T. Bohra, S. Sompura, K. Parekh and P. Raut, "Real-Time Two Way Communication System for Speech and Hearing Impaired Using Computer Vision and Deep Learning," 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2019, pp. 734-739, doi: 10.1109/ICSSIT46314.2019.8987908.

[4] Singha, Joyeeta & Das, Karen. (2013), "Recognition of Indian Sign Language in Live Video," International Journal of Computer Applications. 70. 10.5120/12174-7306.

[5] H. Muthu Mariappan and V. Gomathi, "Real-Time Recognition of Indian Sign Language," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862125.

[6] J. Zhou and J. Hoang, "Real Time Robust Human Detection and Tracking System," in *IEEE Computer Society Conference*, 2005.

[7] C. P. Papageorgiou, M. Oren and T. Poggio, "A General Framework for Object Detection," in *Center for Biological and Computational Learning Artificial Intelligence Laboratory*, Cambridge, 2005.

[8] R. D. Charette and F. Nashashibi, "Real TimeVisual Traffic Lights Recognition Based on Spot Light Detection andAdaptive Traffic Lights Template," in *IEEE*, 2009.

[9] S. K. Nayar, S. A. Nene and M. Hiroshi, "Real-Time 100 Object Recognition System," in *IEEE International Conference*, 1996.

[10] A. Adam, E. Rivlin, S. Ilan and D. Reinitz, "Robust Real-Time Unusual Event Detection Using Multiple Fixed-Location Monitors," in *IEEE*, 2008.

[11] C. Bahlmann, Y. Zhu, R. Vishwanathan, M. Pelkoffer and T. Koehler, "A System for Traffic Sign Detection, Tracking, and Recognition Using Color, Shape, and Motion Information," in *IEEE*, 2005.

[12] M. Betke, E. Haritaoglu and L. S. Davis, "Real-time multiple vehicle detection and tracking from a moving vehicle," *Machine Vision and Applications,* p. 12, 2000.

[13] Q. Chen, N. D. Georganas and E. M. Petriu, "Real-time Vision-based Hand Gesture Recognition Using Haar-like Features," in *IEEE Conference*, 2007.