Ramaiah Institute of Technology

(An Autonomous Institute, Affiliated to VTU)

MSR Nagar, MSRIT post, Bangalore - 54

A Dissertation PBL Report on

Sign Language Recognition

Submitted by

M Sneha	1MS14CS058
Mipsa Patel	1MS14CS148
Tilak S Naik	1MS14CS134
Vibha Karanth	1MS14CS136

Bachelor of Engineering in Computer Science & Engineering

Under the guidance of

Pramod Sunagar
Assistant Professor
Department of Computer Science
Ramaiah Institute of Technology

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

M. S. RAMAIAH INSTITUTE OF TECHNOLOGY
(Autonomous Institute, Affiliated to VTU)
BANGALORE - 560054
www.msrit.edu, 2017

DECLARATION

I Student of seventh semester BE, Dept of Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, hereby declare that the project entitled "Sign Language Recognition", thesis completed and written by me under the guidance of Pramod Sunagar, Dept of CSE, Bangalore for the partial fulfillment of the requirements for the award of the degree of Bachelor of Engineering.

Place: Bangalore
Date: 28th October, 2017

(1MS14CS058 M Sneha)
(1MS14CS148 Mipsa Patel)
(1MS14CS134 Tilak S Naik)
(1MS14CS136 Vibha Karanth)

ACKNOWLEDGEMENT

First and foremost, my utmost gratitude to Pramod Sunagar, Dept of CSE, MSRIT whose sincerity and encouragement we will never forget. He has been our inspiration as we overcame all the obstacles in the completion of this project work.

Dr. Anita Kanavalli, Head of the Department of Computer Science and Engineering, had kind concern and consideration regarding project work and we would like to thank her for continuous support.

We would like to thank our beloved principal Dr. N. V. R. Naidu for his support and encouragement.

This work would not have been possible without the guidance and help of several individuals who in one way or another contributed their valuable assistance in preparation and completion of this study.

We would like to express sincere thanks to all the teaching and non-teaching faculty of CSE Department and my dear friends who helped in all the ways while preparing the report.

Abstract

Sign language is the language used by the deaf and dumb to communicate among themselves and others. Unless the concerned people know the sign language properly, there's a communication barrier between them. In most cases, an interpreter is required to carry out such a conversation. To reduce the dependency of the deaf and dumb on interpreters, we develop a real time sign language detection system that converts the gestures from images and videos into English sentences, using various deep learning techniques.

Contents

1 Introduction

1.1 General Introduction

Sign Language is a vision based language that uses the movements of the hands and facial expressions for communication. Subtle differences among different hand gestures can have a huge impact on the meaning that they convey. Along with this, it is important to prevent the background noise of the image from affecting the gestures being communicated. We consider such problems and challenges and develop an effective sign language recognition system.

1.2 Statement of the Problem

There exist models that convert American Sign Language to sentences. However, most of these techniques are not based on deep learning. These models also cannot handle the nuances of the Indian Sign Language. We use the Indian sign Language and deep neural networks to develop a model that can convert gestures of ISL into English text.

1.3 Objectives of the project

- To develop an Indian Sign Language recognition system.

- To design a deep neural network which can give a better performance and robustness compared to the existing image processing and machine learning models.

1.4 Project deliverables

- An application that obtains frames from sign language communication videos, processes them, and obtains the meaning being conveyed by them in the form of English text.

- Comparison of performance of our model with existing algorithms used for sign language recognition.

1.5 Current Scope

Image and video processing is one of the important fields in Machine Learning and Deep Learning right now. The huge number of images and videos on social media and otherwise has led to an increased interest in the field. The algorithms and techniques developed as a result of this can be used in more impactful fields like helping the deaf and dumb lead more convenient lives.

The hand signs along with the emotions displayed by the individual provide the content and tone of the sentence. A tool to convert their words into a widely used language enhances their ability to communicate.

1.6   Future Scope

Since such applications will always be necessary, research and development of new algorithms to improve the efficiency of interpretation of sign language will be endless. Deep learning has a lot more to offer than what has already been explored. With such possibilities, there will always be scope to develop newer models that can be used in sign language recognition. The models used in this can also be used in other applications such as gesture controlled systems, tracking unusual activity, etc. with slight tweaks. With slight modification, it can also be used for recognition of other sign languages.

Moreover, the current work can be further extended to output speech instead of text. A similar generative model can be used to map speech or text to actions by stitching together pre-recorded videos or actually generate an interface with a skeletal structure. This model can be seamlessly integrated with various applications on different platforms to support high productivity.

2   Literature Survey

2.1   Introduction

There have been a lot of different approaches used to solve the problem of sign language recognition. However, most of these use image processing in MATLAB, upon which Machine Learning algorithms are applied. There has also been some work done using artificial neural networks.

2.2   Related work with the citation of references

[1] deals with the double handed Indian Sign Language. It is captured as a series of images and is processed with the help of MATLAB and then converted to speech and text. However this approach uses image processing techniques which are highly sensitive to lighting conditions.

[2] Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.
    And another one
[3] Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida

sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

2.3 Conclusion of Survey

Deep neural networks have revolutionized how different Machine Learning problems are approached. Using them to solve this problem can give significant results, and that is precisely what this project explores.

3 Project Management Plan

3.1 Schedule of the project

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

4 Software Requirement Specification

4.1 Project Overview

Images of numbers and letters, and videos of words and sentences are fed as input to the model and interpreted text is received as output.

4.2 Hardware

We use a system with NVIDIA® GeForce® 940MX (4 GB DDR3) GPU and 8 GB RAM to perform our experiment. A camera is required to record hand gestures and facial expressions.

4.3 Software Requirements

- Python

- OpenCV

- CUDA

- PyTorch has been used for training and testing the deep neural network.

4.4 Functional Requirements

- Take image and video inputs of sign language conversation.

- Extract frames from videos for training.

- A training algorithm to reduce train-test error and generalize the model to make it invariant to lighting conditions.

- Convert the received output classes to letters, words or sentences.

5 Design

5.1 Introduction

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

5.2 Architecture Design

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

5.3 Data Flow Diagram

Donec et nisl id sapien blandit mattis. Aenean dictum odio sit amet risus. Morbi purus. Nulla a est sit amet purus venenatis iaculis. Vivamus viverra purus vel magna. Donec in justo sed odio malesuada dapibus. Nunc ultrices aliquam nunc. Vivamus facilisis pellentesque velit. Nulla nunc velit, vulputate dapibus, vulputate id, mattis ac, justo. Nam mattis elit dapibus purus. Quisque enim risus, congue non, elementum ut, mattis quis, sem. Quisque elit.

5.4 Conclusion

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

References

[1] Kusumika Krori Dutta, Satheesh Kumar Raju K, Anil Kumar G S, and Sunny Arokia Swamy B. Double handed indian sign language to speech and text. In *Proceedings of the 2015 Third International Conference on Image Information Processing (ICIIP)*, ICIIP '15, pages 374–377, Washington, DC, USA, 2015. IEEE Computer Society.

[2] K. Dabre and S. Dholay. Machine learning model for sign language interpretation using webcam images. In *2014 International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA)*, pages 317–321, April 2014.

[3] K. Dixit and A. S. Jalal. Automatic indian sign language recognition system. In *2013 3rd IEEE International Advance Computing Conference (IACC)*, pages 883–887, Feb 2013.