Real time Indian Sign Language (ISL) Recognition Using YOLOv3

A

SYNOPSIS

SUBMITTED TO THE

SHETH L.U.J AND SIR M.V. COLLEGE

FOR THE DEGREE

OF

BACHLEOR OF COMPUTER SCIENCE

IN

COMPUTER SCIENCE

BY

Subhashish Ramesh Nabajja

(Seat No. S091)

UNDER THE GUIDANCE OF

Prof. _____

DEPARTMENT OF COMPUTER SCIENCE.

SHETH L.U.J AND SIR M.V. COLLEGE,

DR. S. RADHAKRISHNAN MARG, ANDHERI EAST, MUMBAI -400069

Year 2023.

1. Title : **Real time Indian Sign Language (ISL) Recognition Using YOLOv3**
2. Name of student : **Subhashish Ramesh Nabajja**
3. Seat no. : **S091**
4. Subject : **Computer Science**
5. Guide name : _____

Signature of student **Signature and seal of guide**

TABLE OF CONTENTS

INTRODUCTION

A sign language is a language that combines hand shapes, orientation and movement of the hands, arms, as well as body, and facial expression to fluidly express a speaker's thought. Rather than using acoustically transmitted sound patterns, sign languages use visually transmitted sign patterns to convey meaning. The most prevalent sensory impairment in people today is hearing loss. According to WHO estimates, 6.3% of the population in India has a significant hearing impairment, which affects about 63 million individuals. Currently, 291 people out of every 10,000 people in the population have severe to profound hearing loss, according to the NSSO survey (NSSO, 2001)[1]. Children ranging in age from 0 to 14 make up a sizable portion of this group. A significant loss in physical and economic productivity results from the enormous proportion of young Indians who are hearing-impaired. Milder forms of hearing loss and unilateral (one-sided) hearing loss affect an even greater portion of the general public.

A sign language is made up of distinct facial expressions, hand movements, and body and hand orientations. Sign is made up of three major components: Manual features involving gestures made with the hands (employing hand shape and motion to convey meaning), Non-Manual Features (NMFs) such as facial expressions or body posture, which can both form part of a sign or modify the meaning of a manual sign, and Finger spelling, where words are spelled out gesturally in the local verbal language. Naturally, this is an oversimplification; sign language is as complex as any spoken language; each sign language has thousands of signs, each differing from the next by minor differences in hand shape, motion, position, non-manual features, or context.The deaf utilise these gestures to communicate their ideas. But most of the hearing population never tries to learn these gestures, therefore this creates a communication gap between the deaf-dumb people and hearing people. Usually deaf people seek help from sign language interpreters for translating their thoughts to normal people, however this system is very expensive and constrained. Therefore there is a need for a sign language recognition (SLR) system, such a system can automatically recognize gestures, hence minimising the gap in communication.

Worldwide, there are numerous sign languages. The spoken culture and languages of a region determine the sign language that is used there. Indian deaf people communicate using Indian sign language (ISL). Indian Sign Language (ISL) includes both word-level gestures and fingerspelling. Fingerspelling is a method of forming words using letter-by-letter coding. Letter-by-letter signing can be used to express words for which no signs exist, words for which the signer does not know the

gestures, or to emphasise or clarify a specific word. Thus, recognition of fingerspelling is critical in sign language recognition (SLR) system. In Indian sign language, fingerspelling consists of both static and dynamic gestures formed by two hands with arbitrarily complicated shapes.

Most sign language recognition (SLR) system can successfully recognize static signs, but recognizing signs that use non manual features are still an active area of research. Most researchers in this field focus on the recognition of American Sign Language (ASL) because most ASL signs are single handed and thus have a lower level of complexity. Another appealing feature is that ASL already has a standard database that can be used.

Sign language recognition is primarily divided into two approaches: glove-based and vision-based. Signers in the first category must wear a sensor glove or a coloured glove. Wearing the glove during processing simplifies the task of segmentation. The disadvantage of this approach is that the signer must wear the sensor hardware as well as the glove during system operation. The vision-based approach detects and tracks hand signs as well as the signer's facial expressions using image processing algorithms. This method is simpler for the signer because no additional hardware is required. However, there are accuracy issues with image processing algorithms that have yet to be addressed. In vision-based sign language recognition, there are two approaches: 3D model-based and appearance-based. 3D model-based methods make use of 3D information from key body parts. Several important parameters, such as palm position, joint angles, and so on, can be obtained using this information. This method employs either volumetric or skeletal models, or a combination of the two. Volumetric approaches are better suited in the computer animation industry and for computer vision applications.This approach is very computationally intensive, and systems for live analysis are still in the works. Images or videos are used as inputs for appearance-based systems. They make direct inferences from these films and pictures. They do not employ a spatial model of the body. Using a template database, the parameters are taken directly from the pictures or videos. Deformable 2D models of human body parts, particularly the hands, are included in certain templates. Deformable templates are collections of points on an item's outline that serve as interpolation nodes for an approximate outline of the object. These template-based models are mostly used for hand-tracking, but could also be used for simple gesture classification[2]

LITERATURE REVIEW

The literature review is an important part of this research since it helps us grasp the body of information already known about the topic and the prior research that has been done on it by various researchers. Learning about the approaches, the issues they encountered, and how to solve them is also beneficial.

P. Subha Rajam and Dr. G. Balakrishnan[3] developed a sign language recognition system that implemented image processing techniques and can recognize 32 combinations of binary number signs. The system achieved an accuracy of 96.87%. The images were loaded dynamically at run time. Similarly, Divya Deora and Nikesh Bajaj[4] developed a sign language recognition system that uses PCA (Principal Component Analysis) to recognize static signs of Indian Sign Language (ISL). The database constituted 510 images of signs which include 25 alphabets and 9 numbers. Segmentation was done using Red and Blue coloured gloves. The system achieved an accuracy of 94%.

An artificial neural network based sign language recognition system was proposed by Adithya V., Vinod P. R and Usha Gopalakrishnan[5] that can recognize signs of Indian Sign Language that include 26 English alphabets and numerals from 0-9. The system was trained on 360 images with 10 signs of each of the 36 signs, and tested with 180 images with 5 signs of 36 signs. The system achieved an average accuracy rate of 91.11%.

Anuja V. Nair and Bindu V.[2] highlights the limitations of most sign language recognition systems. Most systems can only recognize static hand gestures and the inputs are also constrained, making the system signer-dependent. Research remains to be done on the topic of recognising signs that involve motion. Research works have focused mainly on the recognition of static signs of ISL from images or video sequences that have been recorded under controlled conditions.

In [6] the author proposed a sign language recognition system which works on real-time conditions and makes use of a combinational feature vector with a multiclass support vector machine (MSVM) classifier. The system was able to accurately identify manual signs which comprises hand gestures of isolated signs. The system was trained on 120 images of signs from Indian Sign Language (ISL) of 12 different people, and was tested on 600 images. The system achieved a success rate of 96.23%.

The system proposed by J.Rekha, J.Bhattacharya and S.Majumder in [7] can recognize dynamic signs of Indian Sign Language (ISL). The dataset consisted of 230 training static images and 60 dynamic alphabet videos. The segmentation and

detection of hand is done using skin colour model, which is followed by hand feature extraction using Principal Curvature based Region Detector and 2-D Wavelet Packet Decomposition. The extracted features are converted into appropriate feature vectors. Multiclass non linear support vector machines (SVM) are used for the classification of each hand gesture. The system can successfully recognize both static and dynamic gestures with a success rate of 86.3%.

The system proposed in [8] follows a vision-based gesture recognition system to recognize static, dynamic and finger spelling words of the Indian Sign Language (ISL). The system can recognize gestures in real time. Co-articulation detection between two static gestures, and dynamic gestures and between two dynamic gestures are addressed using a gradient of acceleration approach and obtained 100% accuracy with test dataset. When compared to other systems, the system was proven to be superior.

OBJECTIVES

1. **Real time :** Build a real time sign language recognition system that can recognize static as well dynamic signs.
2. **Scalable and Extendable** : The system should be easily extended to recognize sign languages of different regions.
3. **Detection of gestures with non manual features** : The system should be able to recognize gestures with non manual features i.e facial expression, hand movements, etc.

SCOPE OF RESEARCH

The purpose of this research is to develop a system that uses YOLO model to recognize signs of Indian Sign Language (ISL) from a live video stream. The system must be capable of recognizing static as well as dynamic signs of the Indian Sign Language (ISL).

The research will be conducted over a period of 1 year. The research design is divided into 3 phases.

Phase 1 : Literature Review
Phase 2 : Dataset generation and collection
Phase 3 : Training the model
Phase 4 : Testing the model
Phase 5 : Conclusion and Report writing

PROPOSED METHODOLOGY

The proposed methodology for building a system for sign language recognition (SLR) using YOLO will be first collecting / creating dataset for training and testing our model. The dataset used will include static sign gestures and also dynamic gesture of the Indian Sign Language (ISL).

Once the model has been trained, it will be used for recognizing signs from a real time video stream. A sophisticated web app will be created that will output the meaning of the signs recognized from the image sequences.

REFERENCES

1. National Health Mission (nhm.gov.in)
2. V.Nair, Anuja & V, Bindu. (2013). A Review on Indian Sign Language Recognition. International Journal of Computer Applications. 73. 33-38. 10.5120/13037-0260.
3. Rajam, P. Subha; Balakrishnan, G. (2011). *[IEEE 2011 IEEE 13th International Conference on Communication Technology (ICCT) - Jinan, China (2011.09.25-2011.09.28)] 2011 IEEE 13th International Conference on Communication Technology - Real time Indian Sign Language Recognition System to aid deaf-dumb people. , (), 737–742.* doi:10.1109/ICCT.2011.6157974
4. Deora, D., & Bajaj, N. (2012). Indian sign language recognition. 2012 1st International Conference on Emerging Technology Trends in Electronics, Communication & Networking. doi:10.1109/et2ecn.2012.6470093
5. Adithya, V., Vinod, P. R., & Gopalakrishnan, U. (2013). Artificial neural network based method for Indian sign language recognition. 2013 IEEE CONFERENCE ON INFORMATION AND COMMUNICATION TECHNOLOGIES. doi:10.1109/cict.2013.6558259
6. Dixit, K.; Jalal, A. S. (2013). *[IEEE 2013 3rd IEEE International Advanced Computing Conference (IACC 2013) - Ghaziabad (2013.2.22-2013.2.23)] 2013 3rd IEEE International Advance Computing Conference (IACC) - Automatic Indian Sign Language recognition system. , (), 883–887.* doi:10.1109/IAdCC.2013.6514343
7. Rekha, J.; Bhattacharya, J.; Majumder, S. (2011). *[IEEE Computing (TISC) - Chennai, India (2011.12.8-2011.12.9)] 3rd International Conference on Trendz in Information Sciences & Computing (TISC2011) - Shape, texture and local movement hand gesture features for Indian Sign Language recognition. , (), 30–35.* doi:10.1109/TISC.2011.6169079
8. Nanivadekar, Purva A.; Kulkarni, Vaishali (2014). *[IEEE 2014 International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA) - Mumbai, Maharashtra, India (2014.4.4-2014.4.5)] 2014 International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA) - Indian Sign Language Recognition: Database creation, Hand tracking and Segmentation. , (), 358–363.* doi:10.1109/CSCITA.2014.6839287
9. A Signer Independent Sign Language Recognition with Co-articulation Elimination from Live Videos: An Indian Scenario