A
Synopsis
on

Indian Sign Language Gesture Classification as Single Handed Gestures

submitted in partial fulfillment of the requirements
for completion of AI LAB
of

TY COMP
in
Computer Engineering
by

Harshal Nandavane
MIS NO. 112103095

Ujjwal Pathade
MIS NO. 112103101

Mr. Suraj Sawant
Professor
Department of Computer Engineering

Department of Computer Engineering,
COEP Technological University (COEP Tech)
(A Unitary Public University of Govt. of Maharashtra)
Shivajinagar, Pune-411005,Maharashtra, INDIA

December, 2023

1 Introduction

Hand gestures are a vital form of communication for the hearing impaired, and there's a critical need for systems that can recognize and convert these gestures into meaningful communication to enhance their quality of life. While American Sign Language (ASL), British Sign Language (BSL), and Chinese Sign Language have received substantial attention in gesture recognition projects, Indian Sign Language (ISL) has been relatively overlooked. ISL differs significantly from other sign languages in terms of grammar and structure, utilizing both hands and dynamic gestures, but there's a scarcity of datasets for ISL gestures.

To address this gap, there's a strong need to develop an efficient system capable of accurately interpreting ISL gestures. Such a system not only benefits the hearing impaired in the Indian community but also opens up possibilities for Human-Computer Interaction. To manage the complexity of ISL gestures, they can be classified as single-handed or double-handed, reducing recognition challenges.

2 Motivation

The primary motivation is to promote inclusivity and accessibility in communication. Millions of Deaf and hard-of-hearing individuals around the world rely on ASL as their primary means of communication. This project aims to break down barriers, ensuring that they have equal access to technology and can effectively communicate in a predominantly hearing world.

This project seeks to empower the Deaf community by providing them with the tools to express themselves, access information, and engage in various aspects of daily life independently. It ultimately aims to enhance their quality of life and increase their participation in society.

ASL hand recognition technology can serve as a vital communication support tool in critical areas such as healthcare, emergency services, and customer interactions. It ensures that Deaf individuals can effectively communicate their needs and preferences in various situations.

3 Literature Review

Year	Author Name	Algorithm/Technique used	Remarks
2015	Akanksha Singh, Saloni Arora, Pushkar Shukla, Ankush Mitta	Histogram of Gradients (HOG) features and geometric descriptors using KNN and SVM classifiers were tried on a dataset consisting of images of all 26 English alphabets present in the ISL under variable background.	HOG features when classified with Support Vector Machine were found to be the most efficient approach resulting in an accuracy of 94.23%.
2017	Yogeshwar I. Rokade, Prashant M. Jadav	Firstly, segmentation phase is performed based on the skin colour so as to detect the shape of the sign. The detected region is then transformed into binary image. Later, the Euclidean distance transformation is applied on the obtained binary image. Row and column projection is applied on the distance transformed image. For feature extraction central moments along with HU's moments are used. For classification, neural network and SVM are used.	For feature extraction Central moments and HU moments are used. Artificial neural network is used to classify the sign which gives average accuracy of 94.37% and SVM classifier for the same gives' accuracy of 92.12%.
2022	Shagun Katoch, Varsha Singh, Uma Shanker Tiwary	Segmentation is done based on skin colour as well as background subtraction. SURF (Speeded Up Robust Features) features have been extracted from the images and histograms are generated to map the signs with corresponding labels. The Support Vector Machine (SVM) and Convolutional Neural Networks (CNN) are used for classification. An interactive Graphical User Interface (GUI) is also developed for easy access.	It is achieved by constructing a custom data set, making the system invariant to rotation and solving the background dependency problem. The system is successfully trained on all 36 ISL static alphabets and digits with an accuracy of 99%.

4 Research Gaps

I. Limited Datasets and Resources: Many research papers highlighted the scarcity of comprehensive ISL datasets, making it challenging to develop and evaluate ISL recognition systems. There's a need for larger and more diverse datasets to train and test ISL models effectively.

II. Regional Variation: The linguistic and cultural diversity in India leads to regional variations in ISL. Research often struggled to address these variations comprehensively. Creating standardized models that account for regional differences remains a challenge

III. Gesture Variability: Deaf individuals may use different signing styles or exhibit natural variations in their gestures. Accounting for this variability and ensuring robust recognition across diverse signing styles is a gap in ISL research.

IV. Signer-Independent: Achieving signer-independent recognition, where the system can understand signs from any signer, is a significant research challenge.

V. Ambiguity and Homophony: Similar signs or gestures with different meanings (homophony) and multiple signs representing a single concept (ambiguity) are challenges in ISL recognition. Disambiguating signs and improving the recognition of contextually relevant signs is an ongoing research gap.

5 Problem Statement and Objectives

I. The absence of a standardized and widely recognized Indian Sign Language (ISL) system compounds the daily difficulties faced by Deaf individuals in accessing education, healthcare, employment opportunities, and social inclusion.

   Understanding the precise meaning of deaf and dumb peo- ple's symbolic gestures and converting it into understand- able language

6 Methodology

image segmentation and binarisation The process of image segmentation and binarization in the context of converting colored images to binary format using RGB thresholding techniques involves comparing the differences between the Red and Blue color components. Specifically, a difference greater than 25 is used to identify skin pixels, while a difference greater than 20 is considered significant for Red dominance. After this initial thresholding, morphological operations, such as dilation and closing, are applied to the binary image to further eliminate noise. Additionally, the largest connected object in the binary image, typically representing alphabets formed by two hands in sign language (ISL), is extracted, effectively removing any remaining unwanted noise. The result is a clean binary representation of the image.

Geometric descriptors used for classifying gestures as single or double-handed include Solidity, representing the area-to-convex area ratio, which gauges compactness; the Major to Minor Axis Ratio, indicating the image's spread; Eccentricity, characterizing object roundness; the Area-to-Bounding Box Ratio, measuring object fit within a bounding box; and the Equivalent Diameter-to-Minor Axis Ratio, which assesses object size and shape. These descriptors offer key information for distinguishing between single and double-handed gestures in binary images.

HOG features are employed to detect object shapes by capturing local orientations and intensity distributions. Images are divided into cells, and the collective orientation of pixels in each cell forms a histogram of gradients. To enhance illumination invariance, cell intensities are normalized by block illumination. Blocks are made by sliding cells, allowing for overlap. The Histogram of Oriented Gradient summarizes gradient magnitudes in various directions for all cells and blocks. In the context of binary images, HOG is applied after converting them to grayscale format. Each image yields a feature space with 81 features, generated from 9-bin histograms per cell and 9 rectangular cells.

7 Hardware and Software Requirements

The successful implementation of AI algorithms, Support Vector Machine(SVM),Convolutional Neural Network (CNN), requires a standard computer system equipped with a multi-core processor(min i3) and ample RAM. In the event of potential neural network integration, access to a GPU (Graphics Processing Unit) is recommended for expedited model training. The software prerequisites encompass commonly used programming libraries and tools, such as Python, Indian Sig Language Datasets, and OpenCV development and testing.

8 Conclusions

Hand detection project employed a Random Forest method, deviating from the commonly utilized CNN and SVM approaches found in existing research papers. The choice of Random Forest demonstrates the versatility and effectiveness of alternative machine learning techniques in the domain of hand detection. Through this project, valuable insights have been gained into the potential of diverse methodologies for achieving accurate and efficient hand detection, contributing to the advancement of the field.