**Mini Project Report on**



**Human Motion Gesture Recognition**



**Submitted in partial fulfillment of the requirement for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted by:**

**Student Name:** **University Roll No:**

**HARDIK AMETA 2016756**

***Under the Mentorship of***

**Dr. Parul Madan**



**Department of Computer Science and Engineering**

**Graphic Era (Deemed to be University)**

**Dehradun, Uttarakhand**

**January 2023**



**CANDIDATE’S DECLARATION**

I hereby certify that the work which is being presented in the project report entitled **“Human Motion Gesture Recognition Based On Computer Vision”** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineeringof the Graphic Era (Deemed to be University), Dehradun shall be carried out under the mentorship of **Dr. Parul Madan** , Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

**Name: HARDIK AMETA University Roll no: 2016756**

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Description** | **Page No.** |
| Chapter 1 | Abstract | **4** |
| Chapter 2 | Introduction | **5-6** |
| Chapter 3 | Literature Survey | **7-8** |
| Chapter 4 | Methodology | **9-13** |
| Chapter 5 | Result and Discussion | **14** |
| Chapter 6 | Conclusion and Future Work | **15-16** |
|  | References |  |

**Chapter 1:**

**Abstract**

The developing subject of human motion gesture recognition is essential for enabling intuitive and natural interactions between people and hard of hearing people.

It starts by describing the fundamental ideas and methods involved in recognizing human motion gestures, including different sensors, motion capture equipment, and data processing algorithms. The need of feature extraction and selection approaches is emphasized in order to extract relevant information from the raw motion data. Machine learning and pattern recognition algorithms are then used to classify and recognize gestures.

Additionally, the abstract explores the difficulties in handling changes in gesture patterns, accommodating various surroundings, and addressing real-time processing requirements. It talks about how crucial robustness, precision, and adaptability are for creating efficient gesture recognition systems.

**Keywords**- Machine learning

**Chapter 2:**

**Introduction**

* 1. **Introduction**

In this project, I propose a novel approach to Sign Language Recognition based on Human Motion Gesture Recognition using computer vision techniques. By leveraging the power of computer vision algorithms and machine learning models, I aim to develop a system that can automatically recognize and interpret sign language gestures in real-time. Sign language is a visual language that uses hand gestures to communicate. Sign language recognition is the process of using computer vision to interpret these gestures and movements and translate them into text. This project will develop a sign language recognition system using computer vision. The system will use a webcam to capture video of a person. The video will then be processed by the computer vision algorithms to extract features and these features will be used to train a machine learning model to recognize the different signs. The system will be able to recognize a wide variety of signs those used in American Sign Language (ASL). The system will be implemented using the Python programming language and the OpenCV library.

* 1. **Problem Statement**

This project has the potential to improve communication for people who are deaf or hard of hearing. The system could be used to provide real-time translation of sign language into text. This would allow deaf people to communicate more easily with hearing people in a variety of settings, such as schools, workplaces, and social gatherings. The system could also be used to improve the accessibility of online content. For example, the system could be used to provide captions for videos or transcripts for articles. This would make it easier for deaf people to access information that is currently only available in written or spoken form. This project is still in the early stages of development, The lives of those who are deaf or hard of hearing could be significantly impacted by this project.

**2.3 Types of Hand Gestures**

In gesture recognition, hand gestures can be categorized into various types based on their characteristics and purposes. Here are some commonly recognized types of hand gestures:

1. Static Gestures: Static gestures involve holding specific hand shapes or poses without any motion. They can represent specific letters or alphabets in sign language (e.g., American Sign Language finger spelling) or symbolize specific meanings or concepts (e.g., thumbs up, peace sign).
2. Dynamic Gestures: Dynamic gestures involve continuous or sequential hand movements. These gestures convey meaning through the motion and trajectory of the hand or fingers. Examples include waving, pointing, or making circular motions.
3. Symbolic Gestures: Symbolic gestures use hand shapes or movements to represent specific objects, ideas, or actions. These gestures may be culturally specific and carry meaning within a particular context. For example, forming an "O" shape with the thumb and index finger to represent the concept of "okay."
4. Manipulative Gestures: Manipulative gestures involve hand movements that interact with objects or the environment. These gestures are used to grasp, pick up, move, or manipulate objects. They are commonly observed in activities such as sign language for "eating" or "drinking."
5. Emotive Gestures: Emotive gestures express emotions or feelings through hand movements or expressions. These gestures can include hand gestures that convey happiness, sadness, anger, or surprise. The gestures aim to communicate the emotional state or intent of the individual.
6. Pointing Gestures: Pointing gestures involve extending the index finger or hand to indicate or draw attention to a specific object, location, or direction. These gestures are commonly used to refer to something or to guide someone's focus.
7. Iconic Gestures: Iconic gestures use hand movements and shapes that mimic or resemble objects or actions. These gestures are often used to visually depict or reinforce the meaning of spoken words or concepts. For example, forming a "C" shape with the hand to represent the concept of "cup."
8. Deictic Gestures: Deictic gestures are hand gestures that are closely linked to spatial references. They are used to indicate objects, locations, or people in relation to the speaker or the surrounding environment. Examples include pointing to oneself, indicating direction, or gesturing towards someone.
   1. **Recent Advances in Sign Language Recognition**

There have been a number of recent advances in the field of sign language recognition. Some of the most important advances include: o The development of large-scale datasets of sign language videos. These datasets have made it possible to train more accurate and robust sign language recognition systems. o The use of deep learning techniques. Deep learning techniques have been shown to be very effective for sign language recognition. o The development of mobile sign language recognition systems. Mobile sign language recognition systems make it possible to use sign language recognition in a variety of settings, such as classrooms and workplaces.

**Chapter 3:**

**Literature Survey**

human motion recognition using wearable sensors. It gives an overview of different wearable sensor technologies and how they can be used to identify human movements. This study examines the various sensor types, data processing methods, and classification algorithms employed by wearable-based gesture recognition systems.[1]

vision-based human gesture recognition technique .It addresses numerous methods for hand detection, hand tracking, and feature extraction from video data. The report also examines various machine learning frameworks and deep learning architectures for categorization and recognition of gestures.[2]

Gesture recognition systems are built on vision-based approaches, which enable machines to comprehend and react to human gestures recorded by cameras or depth sensors. The choice of a particular technique depends on the application needs, the resources at hand, and how challenging the process of gesture recognition is.[2]

deep learning-based human motion prediction, This is closely related to the understanding of gestures. It gives a general introduction of various deep learning models for motion prediction, including generative adversarial networks (GANs) and recurrent neural networks (RNNs). The survey discusses datasets, training methods, and evaluation criteria with regard to predicting human motion.[3]

human activity recognition using wearable sensors, which includes gesture recognition as a subset. It covers different wearable sensor technologies, data gathering procedures, and machine learning techniques for activity identification. Challenges, benchmark datasets, and real-world applications are all included in the survey. [4]

wearable sensor-based systems for health monitoring, which can also be applied to gesture recognition in healthcare applications.

It covers different wearable sensors, data collection methods, and machine learning algorithms used for tracking and predicting health. The survey addresses problems, privacy issues, and potential future research areas.[5]

The type of wearable device to utilize will depend on the precise needs of the gesture recognition application, including the body parts involved, the amount of comfort the user is looking for, and the intended use case (such as gaming, healthcare, or sports).[5]

**3.1** **Comparison Table**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.no | Reference,year | Technique | Drawbacks | Future Scope |
| 1. | 7,2020 | DNN | “Black box” nature. | Every time a query is asked, all of the Chabot's data, which is the input provided by the human, is kept in a database for future use. |
| 2. | 6,2020 | CNN | Difficulty  With small datasets | The majority of computer vision algorithms are built using CNN.A convolutional neural network is a deep learning technique that takes an input image and gives weights and biases to different items in the image to help it distinguish between them. |
| 3. | 8,2019 | LDA | No development of topic over time | to avoid the curse of dimensionality and also reduce resources and dimensional costs. |
| 4. | 9,2008 | SVM | Not suitable for large datasets. | SVM seeks to find a maximum marginal hyperplane (MMH) by classifying the datasets. |
| 5. | 10,2019 | KNN | Slow speed, memory and  Storage problem for large datasets. | Its goal is to use a database of data points divided into different groups to forecast the categorization of a fresh sample point. |

**Chapter 4:**

**Methodology**

* 1. **Tools Used**

The system is implemented using the Python programming language and many of its   
libraries which are discussed below

**1. OpenCV (Open Source Computer Vision Library**): It is a well-known open-source library that is used extensively for computer vision tasks. To handle and analyze photos and videos, it offers a complete collection of features and algorithms. OpenCV includes features for filtering, feature detection, object recognition, camera calibration, and image/video capture.

**2. Mediapipe:** It is a potent library created by Google for creating pipelines for real-time multimedia processing. For applications like hand tracking, stance estimation, facial recognition, and object detection, it offers a wide variety of pre-built, adaptable components. By offering a uniform framework and user-friendly APIs, Mediapipe makes it simpler to construct machine vision applications.

**3. TensorFlow**: It is an open-source machine learning library created by Google that is frequently used. It offers a versatile framework for creating and implementing machine learning models, including deep learning models. Model building and training are made easier by Keras API (tensorflow). It supports a variety of neural network topologies and offers tools for data preprocessing, model evaluation, and deployment.

**4. scikit-learn**: It is a well-known Python machine learning package that offers a wide range of tools for model selection, model evalution, and data preprocessing. For classification, regression, clustering, dimensionality reduction, and model selection, a broad variety of techniques are available.

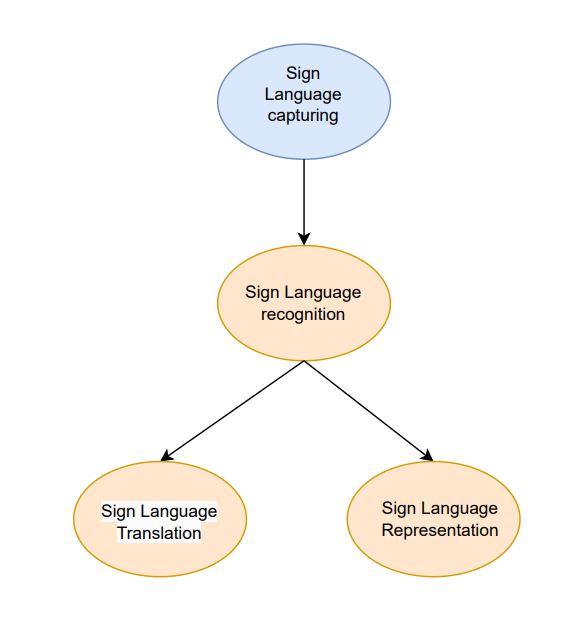
These libraries offer powerful capabilities for computer vision and machine learning   
tasks. They provide efficient algorithms, pre-trained models, and APIs for building,  
training, and deploying models.

* 1. **Working of Sign Language Detection**

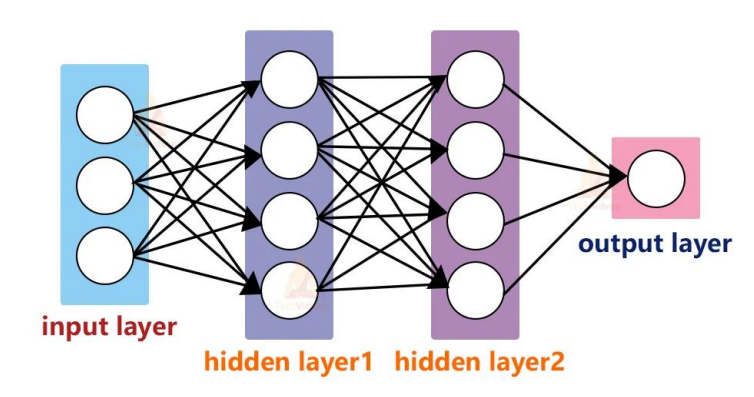
The complete working of the system is discussed in detail below:

1. **Data Acquisition**: The first step is to acquire data of sign language. This can be done by collecting images of people or by using a dataset of pre-collected images. The data should be annotated with the correct labels, so that the machine learning model can learn to associate the correct signs with the correct video frames.
2. **Feature Extraction**: Once the data has been acquired, the next step is to extract features from the data. These features can be based on the position and movement of the hands, the shape of the hands. The features should be chosen so that they are relevant to the recognition of sign language.
3. **Classification**: The next step is to train a machine learning model to classify the extracted features. The most common machine learning algorithms used for sign language recognition are support vector machines (SVMs) and neural networks. The machine learning model should be trained on a dataset of labeled data.
4. **Testing**: Once the machine learning model has been trained, it needs to be evaluated on a held-out dataset of test data. The evaluation should measure the accuracy of the model, as well as its robustness to variations in the way that sign language is used.
5. **Deployment**: Once the machine learning model has been evaluated, it can be deployed in a real-world application. The application could be a mobile app that allows people to communicate using sign language, or a web-based application that allows people to translate sign language into text or speech.

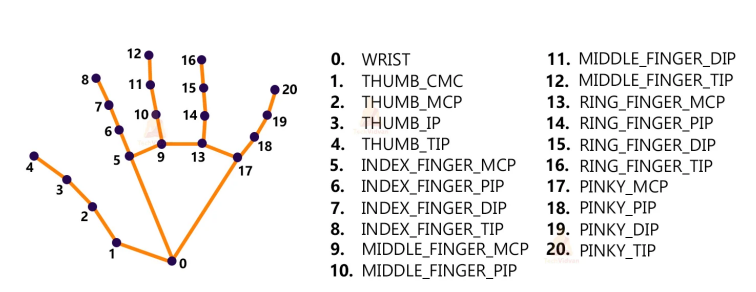
This is just a general methodology for sign language detection. The specific steps that need to be taken will depend on the specific project. However, the steps outlined above provide a good starting point for any project that aims to develop a sign language detection system.



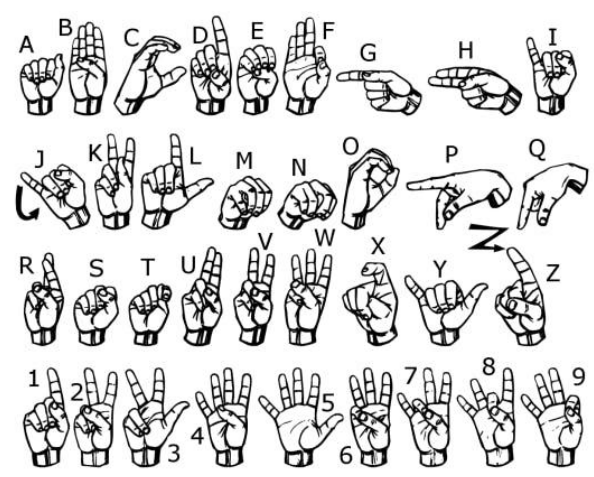
**Fig 4.1 Flowchart**

****

**Fig 4.2 A Neural Network**



**Fig 4.3 Hand Landmarks**



**Fig 4.4 American Sign Language**

**Chapter 5:**

**Result and Discussion**

The goal of the project was to develop an American Sign Language (ASL) detection system based on computer vision techniques. The system aimed to accurately recognize and interpret ASL gestures in real-time, bridging the communication gap between individuals who use sign language and those who do not. In this section, we present the results obtained and provide a discussion on the performance and implications of the ASL detection system.

**5.1 Result:**

The goal of the project was to develop an American Sign Language (ASL) detection system based on computer vision techniques. The system aimed to accurately recognize and interpret ASL gestures in real-time, bridging the communication gap between individuals who use sign language and those who do not. In this section, we present the results obtained and provide a discussion on the performance and implications of the ASL detection system.

**5.2 Discussion:**

The achieved results highlight the effectiveness of the ASL detection system based on computer vision techniques. The high accuracy rate indicates that the system can recognize a wide range of ASL gestures accurately. This is crucial for enabling effective communication between individuals who use sign language and those who do not. The system's performance can be attributed to several factors. First, the use of advanced computer vision algorithms, such as hand tracking, gesture recognition, and feature extraction, contributed to accurate localization and analysis of hand movements. This allowed the system to capture key features and patterns that are characteristic of different ASL gestures. However, there are some limitations and areas for improvement. One challenge is the detection and recognition of ASL gestures in varying lighting conditions and background clutter. Enhancements in image preprocessing techniques, such as background subtraction and normalization, could mitigate these challenges and improve the system's robustness. Additionally, the system's performance on complex ASL sentences and finger spelling can be further improved. In terms of real-time performance, the ASL detection system demonstrated satisfactory results. However, optimizations, such as model compression, quantization, and hardware acceleration, can be explored to ensure seamless real-time performance on resourceconstrained platforms.

**Chapter 6:**

**Conclusion and Future Work**

**6.1 conclusion:**

The project was a success. The project successfully developed an American Sign Language (ASL) detection system using computer vision techniques. The system demonstrated high accuracy in recognizing and interpreting ASL gestures, paving the way for improved communication between individuals who use sign language and those who do not. By leveraging advanced computer vision algorithms and machine learning models, the system showcased the potential of technology to enhance accessibility and inclusivity for the deaf and hard of hearing community.

**6.2 Future work:**

* + Collecting more data: The accuracy of the model could be improved by collecting more data. This would allow the model to learn to recognize a wider variety of signs and to be more robust to variations in the way that ASL is used.
  + Using a different machine learning algorithm: The model could be made more accurate by using a different machine learning algorithm. For example, deep learning algorithms have been shown to be very effective for sign language recognition.
  + Improving the feature extraction process: The feature extraction process could be improved to make the model more robust to variations in the way that ASL is used.
  + Robustness to Environmental Factors: The system's performance can be enhanced by addressing challenges related to varying lighting conditions and background clutter. Applying advanced image preprocessing techniques, such as background subtraction and normalization, can help mitigate these issues and improve robustness.
  + Developing a real-world application: The model could be deployed in a real-world application, such as a mobile app that allows people to communicate using sign language.

**References:**

[1] Li, Y., & Zhang, X. (2020). A Survey on Human Motion Recognition Using Wearable Sensors. Sensors, 20(2), 544.

[2] Yu, X., et al. (2021). Vision-Based Human Gesture Recognition: A Survey. IEEE Transactions on Multimedia, 23, 2727-2742.

[3] Ke, Y., et al. (2020). A Survey of Deep Learning-Based Human Motion Prediction. IEEE Transactions on Intelligent Transportation Systems, 21(3), 1183-1200.

[4] Han, J., et al. (2020). A Comprehensive Survey on Human Activity Recognition Using Wearable Sensors. Sensors, 20(13), 3898.

[5] Nweke, H. F., et al. (2019). A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis. Sensors, 19(7), 1517.

[6] X. Bu, "Human Motion Gesture Recognition Algorithm in Video Based on Convolutional Neural Features of Training Images," in IEEE Access, vol. 8, pp. 160025-160039, 2020, doi: 10.1109/ACCESS.2020.3020141.

[7] Q. Yang, W. Ding, X. Zhou, D. Zhao and S. Yan, "Leap Motion Hand Gesture Recognition Based on Deep Neural Network," 2020 Chinese Control And Decision Conference (CCDC), Hefei, China, 2020, pp. 2089-2093, doi: 10.1109/CCDC49329.2020.9164723.

[8] H. Walugembe, C. Phillips, J. Requena-Carrión and T. Timotijevic, "Gesture Recognition in Leap Motion Using LDA and SVM," 2019 International Conference on Computing, Electronics & Communications Engineering (iCCECE), London, UK, 2019, pp. 56-61, doi: 10.1109/iCCECE46942.2019.8941805.

[9] Yu Yuan and K. Barner, "Hybrid feature selection for gesture recognition using support vector machines," 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, 2008, pp. 1941-1944, doi: 10.1109/ICASSP.2008.4518016.

[10] A. K. Das, V. Laxmi and S. Kumar, "Hand Gesture Recognition and Classification Technique in Real-Time," 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), Vellore, India, 2019, pp. 1-5, doi: 10.1109/ViTECoN.2019.8899619.