E-COMMERCE DIGITAL VOICE ASSISTANT FOR VISUALLY AND HEARING-IMPAIRED USERS WITH CHATBOT – DVC

Project ID: 23-267

S.Sathurshiga

(IT20114480)

Bachelor of Science (Hons) in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

September 2023

DECLARATION

I declare that this is my own work, and this Thesis does not incorporate without acknowledgement any material previously submitted for a degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the non-exclusive right to reproduce and distribute my Thesis, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: ……. Date : 09.09.2023

The above candidate has carried out this research thesis for the Degree of Bachelor of Science (honors) Information Technology (Specializing in Information Technology) under my supervision.

Signature of the supervisor: ……. Date: 10.09.2023

Acknowledgment

We would like to express our gratitude to Our Supervisor **Mr. Uditha Dharmakeerthi** and Co-Supervisor **Mr. Didula Chamara** for their patient guidance, encouragement & useful critiques of this research work. And I would also like to thank our research panel for their advice & feedback in keeping our progress on the schedule. Finally, we wish to thank our parents for their support & encouragement throughout over side.

List of Abbreviation

Abbreviation	Description
ML	Machine Learning
ASL	American Sign Language
BSL	British Sign Language
AI	Artificial Intelligence
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network

1. INTRODUCTION

1.1 Background

Deaf and hard-of-hearing people use sign language, a visual language, to communicate with each other and the hearing community. It is an advanced language that combines body language, face expressions, and hand gestures to convey meaning. There are numerous distinct sign languages used in various nations and locations, hence sign language is not a universal language. Both American Sign Language (ASL) and British Sign Language (BSL), which are used in different parts of the world, are different from other sign languages. Although the deaf and hard-of-hearing community relies heavily on sign language as a means of communication, it can be difficult for them to access digital content that is primarily intended for hearing people. To address this issue, researchers and developers have been working on creating sign language recognition technology that can translate sign language into written or spoken language in real-time. Sign language recognition technology uses a combination of computer vision and Machine learning algorithms to analyze and interpret the movements of the signer's hands, face, and body, and then converts them into written or spoken language. The development of sign language recognition technology has been ongoing for several decades. In the early 1990s, researchers began exploring the use of computer vision techniques to recognize hand gestures in sign language. However, these early systems were limited by the technology available at the time and were not very accurate. Advances in computer vision and machine learning algorithms in recent years have led to significant improvements in the accuracy and reliability of sign language recognition technology. Researchers have developed systems that can recognize a wide range of sign languages and can even interpret complex grammatical structures and nuances of sign language.

Sign language recognition technology has the potential to revolutionize the way deaf and hard-of-hearing individuals interact with digital systems. It can provide a more accessible and inclusive experience for them and promote equal access to digital systems for all individuals. Incorporating sign language recognition systems into online shopping websites, for example, could make them more accessible and user-friendly for deaf and hard-of-hearing individuals, allowing them to shop online independently and with greater ease. The accessibility of the deaf and hard of hearing community could be considerably improved through sign language recognition on online platforms. It does, however, create technical difficulties that must be resolved through continual study and development. It is expected that as technology develops, sign language recognition will become more accurate and available, enabling accessibility and communication in the digital world.

1.2 Background Literature

"Real-Time Sign Language Recognition Using a Kinect Sensor" [1]

In this study, the authors developed a real-time sign language recognition system using a Microsoft Kinect sensor. They used a combination of computer vision and machine learning algorithms to recognize hand gestures, facial expressions, and body language. The system achieved an accuracy of 92.5% in recognizing 21 commonly used signs in American Sign Language (ASL). The authors concluded that the system had the potential to improve accessibility and communication for deaf and hard-of-hearing individuals.

"Sign Language Recognition Using Recurrent Neural Networks with Long Short-Term Memory" [2]

In this study, the authors proposed a sign language recognition system based on recurrent neural networks (RNNs) with long short-term memory (LSTM). They used a dataset of ASL signs and achieved an accuracy of 89.2% in recognizing 10 signs. The authors concluded that the system could be improved by using larger datasets and by incorporating other features such as facial expressions and body language.

"Sign Language Recognition Based on Dynamic Features and Deep Learning" [3]

In this study, the authors proposed a sign language recognition system based on dynamic features and deep learning. They used a dataset of Chinese Sign Language (CSL) signs and achieved an accuracy of 94.1% in recognizing 600 signs. The authors concluded that their system outperformed other state-of-the-art sign language recognition systems and had the potential to be used in real-world applications.

"Sign Language Recognition Using 3D Convolutional Neural Networks" [4]

In this study, the authors proposed a sign language recognition system based on 3D convolutional neural networks (CNNs). They used a dataset of Indian Sign Language (ISL) signs and achieved an accuracy of

94.23% in recognizing 24 signs. The authors concluded that their system had the potential to be used in real-world applications and could be extended to recognize more signs.

"A Wearable System for Real-Time Sign Language Recognition Using Surface Electromyography" [5]

In this study, the authors proposed a wearable sign language recognition system using surface electromyography (sEMG). They used a dataset of ASL signs and achieved an accuracy of 93.6% in recognizing 26 signs. The authors concluded that their system was more practical and comfortable to use than other sign language recognition systems and had the potential to be used in real-world applications. One promising approach to screen navigation using sign language involves the use of computer vision and machine learning techniques In a study by [12] , the authors proposed a system that uses a deep learning-based approach to recognize sign language gestures and translate them into commands for screen navigation. The system achieved an accuracy rate of over 90% in recognizing sign language gestures, demonstrating its potential for improving the accessibility of digital platforms for individuals with hearing and speech disabilities. Another study by [13] proposed a screen navigation system that uses a wearable device to capture and interpret sign language gestures. The authors developed a wearable device that uses sensors to capture hand movements and a machine learning algorithm to translate these movements into screen navigation commands. The system achieved an accuracy rate of over 80% in recognizing sign language gestures, highlighting its potential for improving accessibility.

The literature survey highlights some recent studies and developments in sign language recognition technology. These studies have shown that sign language recognition technology has the potential to improve accessibility and inclusivity for deaf and hard-of-hearing individuals. The studies have used various techniques such as computer vision, machine learning, and wearable technology to recognize different sign languages. While the accuracy of sign language recognition technology has improved over the years, there is still scope for further research and development in this field. Future studies could focus on improving the accuracy of sign language recognition systems, expanding the datasets used for training and testing, and developing more practical and comfortable wearable systems for real-world applications.

1.3 Research Gap

There aren't enough research or reviews of the usefulness and effectiveness of implementing sign language into e-commerce platforms. Although the advantages of sign language recognition for increasing accessibility and inclusion are well known, additional research is required to determine how well this approach will support individuals with diverse abilities in an e-commerce environment. The complexity of ecommerce-related communication, such as product descriptions, prices, and shipping information, is difficult for sign language recognition systems to handle. Additionally, there can be difficulties in ensuring the consistency and correctness of sign language translations, as well as problems with user acceptance and adoption. Therefore, more research was required to determine whether sign language can be effectively integrated into e-commerce platforms, as well as to identify and address any potential obstacles or difficulties. A more accessible and inclusive ecommerce platform design and execution might be aided by such research, which would ultimately enhance the online buying experience for all users.

1.4 Research Problem

The main research problem is **"How to complete online shopping with sign language gestures?"** for visually impaired people. Although there hasn't been much research on how well sign language recognition performs in online stores, it is a good idea to use it to increase accessibility and inclusivity of online buying. Because we are unable to improve online purchasing for everyone, including those with diverse capacities, this information gap is a concern. We need to research whether adding sign language to internet commerce is useful and practical. We must also identify and address any issues that may arise during this process. No matter their preferred method of communication, this data can help us create online stores that are simpler to use for everyone.

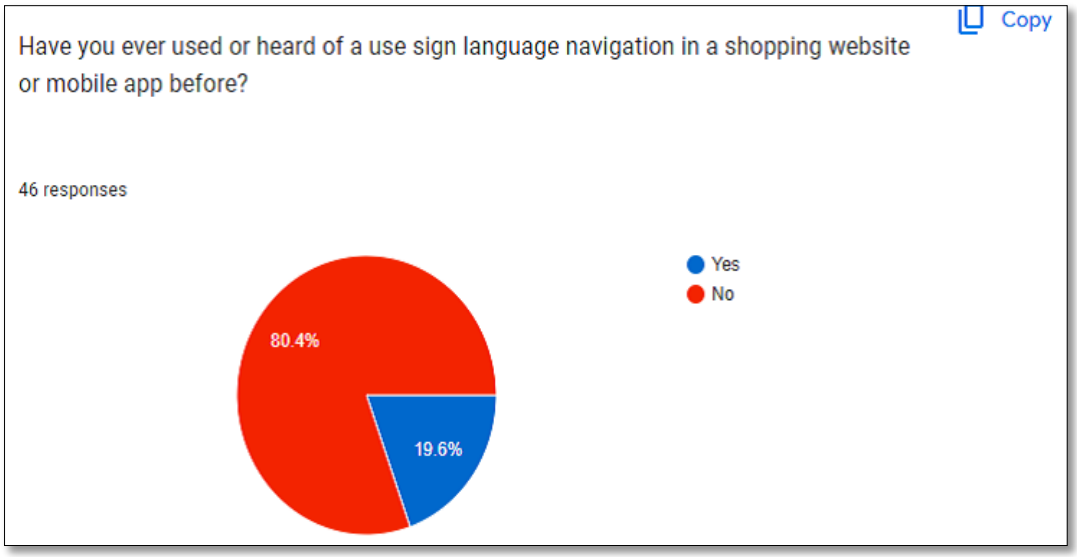


Figure 1: Summary of responses for Have you ever heard about sign language?

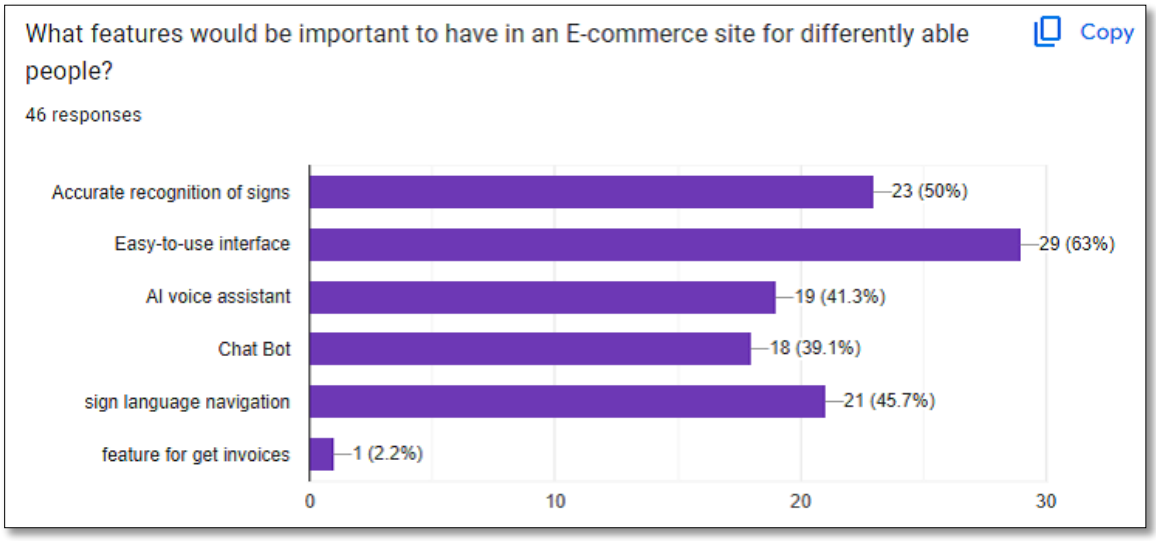


Figure 2: Summary of response for what features would be important to have an E-Commerce site for differently abled people?

According to figure 1, more than 80% of the people in our country never knew or heard about sign language used in an E-commerce platform. That is why we chose this component that need to be add in our system as unique one. And also from the 2nd figure we can conclude that sign language navigation can help the visually impaired people.

2. OBJECTIVES

2.1. Main Objective

Main objective of our research project is to deliver a web application including sign language recognition, AI chat bot and voice assistant. In my component I was considering using sign language detection mechanism to enable differently abled individuals to make purchases through gesture recognition, which allows them to interact with digital systems more efficiently. The system uses machine learning and AI algorithms to analyze and interpret the nuances of different signs and gestures. The approach promotes inclusivity and accessibility in the online shopping experience

2.2Specific Objectives

The specific objectives of this workflow is to ensure the accuracy and effectiveness of the sign language recognition system by:

- **Data Collection**: Compiling a large and varied dataset of sign language movements for the purpose of training ML and AI algorithms. To make the system more inclusive, this dataset should include a variety of sign languages, dialects, and gestures.
- **Data Preparation**: The process of purifying and organizing the collected data. This include cleaning up the data, making sure consistency is maintained, and setting it up for efficient training.
- **Model training**: The process of learning machine learning models using deep learning techniques. In order for the models to effectively learn and identify the complexity and nuance of various signals and gestures, this process involves putting the prepared data into them.
- **Test and Validate**: Test and improve to increase the ML model's accuracy in identifying sign language actions in e-commerce.

The workflow focuses on these sub-objectives in order to develop a strong and reliable sign language identification system that improves the accessibility and inclusivity of digital systems, notably in the context of online shopping for people with disabilities.

3. METHODOLOGY

The system method includes the implementation of Convolutional Neural Networks (CNNs) to recognize sign language in two different ways. The first method examines every video frame separately. It separates the video into individual frames, much like slideshows of images. The CNN is then given each frame as input. The CNN gains the ability to identify crucial details in each frame and track how they alter over time. The CNN can identify the sign language gesture being made by examining these features in each frame. The second strategy is rather different. It examines both how things move and how they appear to be.

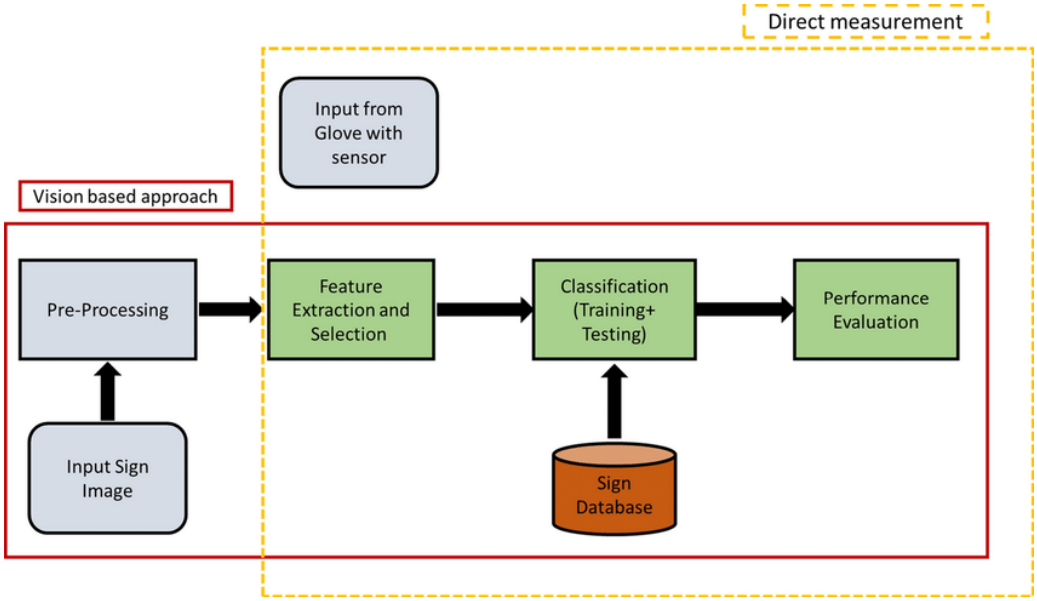


Figure 3:CNN Design

CNNs are especially good for applications involving the analysis of visual data such as sign language recognition since they are built to learn and extract features from photos or videos. In order to recognize diverse sign language motions, a CNN model is trained on a collection of videos. The model uses this information to recognise new gestures after learning to recognize the visual characteristics that set one gesture apart from another.

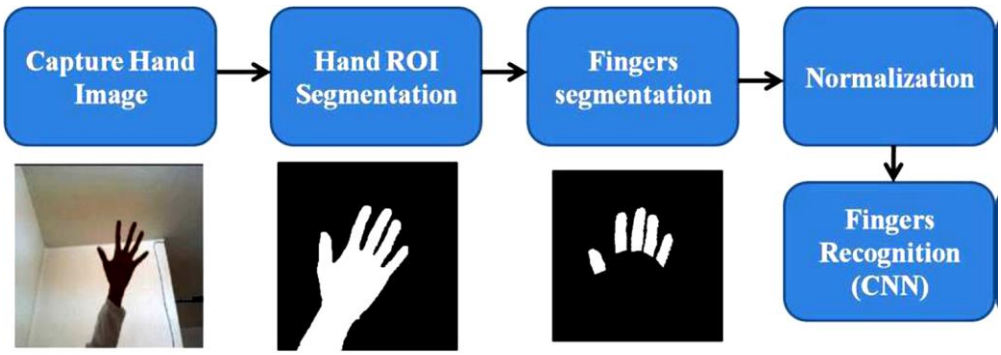


Fig 4: Hand recognition system

3.1 System Overview

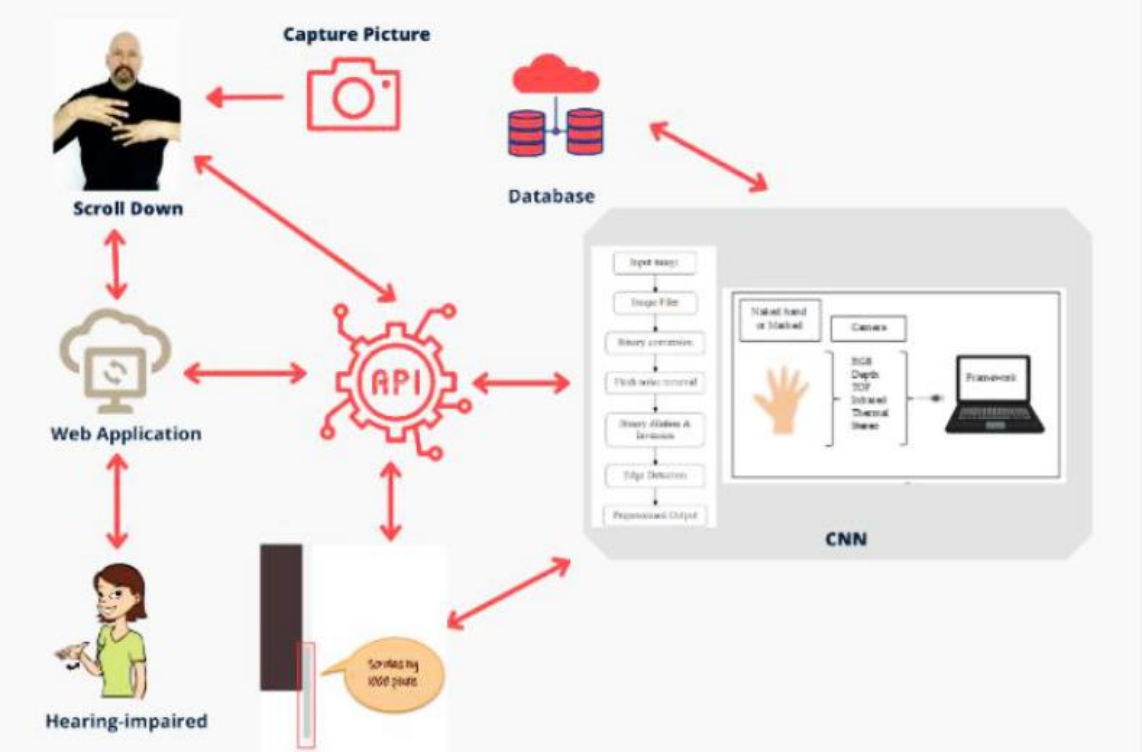


Figure 5: High-Level diagram of system

The E-commerce Platform with Sign Language Recognition is a complex digital system created to improve inclusion and accessibility for people with disabilities, especially those who use sign language to communicate. This system combines modern technologies, such as computer vision and machine learning, to allow users to interact with the platform successfully.

Key Elements:

Front-End Interface: The part of the platform that is visible to users and may be accessed through web browsers since it is a web application. It offers a simple user experience for ordering, exploring products, and purchasing.

Sign Language Recognition: Gesture detection, categorization, and translation, it has separate components.

Database Management: Product listings, user profiles, order histories, and other essential data are stored in databases. It also controls how material is displayed on the platform.

Machine learning model: Skilled at effectively identifying sign language gestures. Its performance is continuously enhanced and optimized to boost recognizing abilities.

Functionalities of the system are users can interact with the application by using sign language gestures that were captured by a webcam, For navigation and interaction, the system decodes these motions and converts them into text or speech commands. Using both sign language and conventional text inputs, users can look up products, read in-depth descriptions, and navigate through categories, Users can manage their shopping carts and wish lists as well as place orders, see order histories, and manage their wish lists, Users can build and manage their own profiles, complete with contact preferences and personal data.

Users can use regular text inputs or camera-captured sign language motions to interact with the site. The gestures are processed and converted into commands by the Sign Language Recognition Module. Through feedback and optimization, the Machine Learning Model continuously improves its comprehension of sign language gestures. Product data, order information, and other pertinent stuff are displayed in the front-end interface. User profiles, product listings, and transaction records are all stored and retrieved in the database. Users receive notifications, order confirmations, and other messages through the communication system. Online payments are processed safely by the payment gateway.

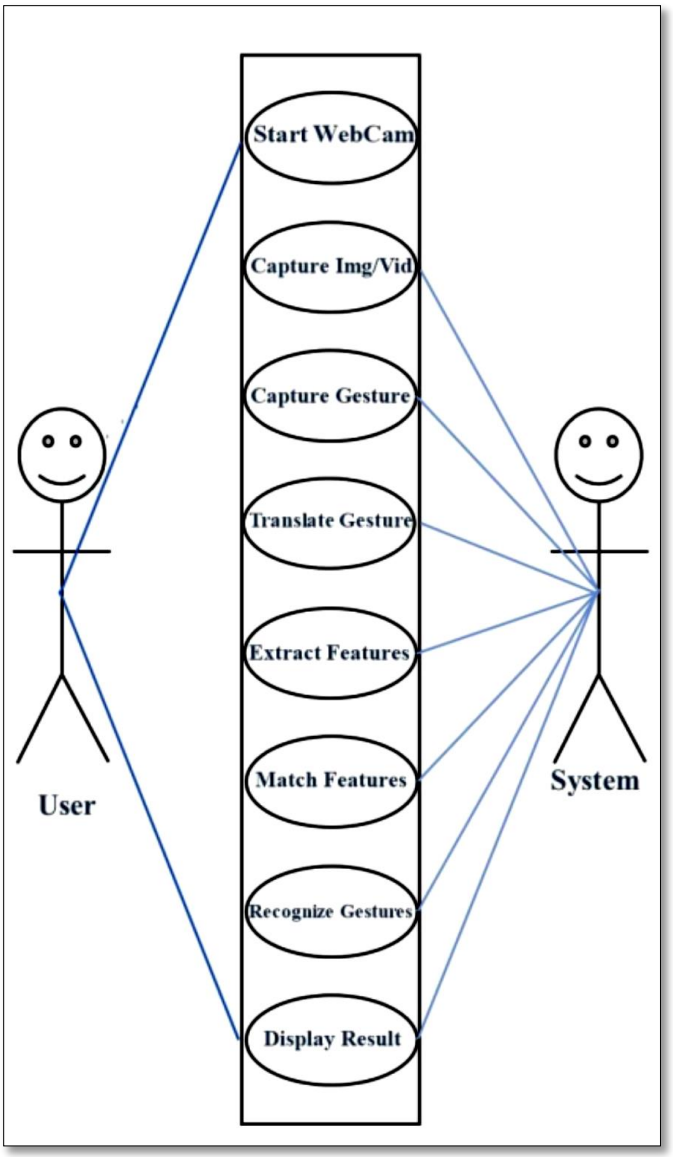


Figure 6: Use Case diagram for Automatic irrigation System

Used datasets.

Data got from Kaggle

3.2 Resource Used

Software Boundaries

- ***IDE***: Pycharm, Visual Studio Code, IntelliJ
- ***Languages***: Python, Java, HTML, CSS, JS, JQuery
- ***Frameworks***: Bootstrap, Springboot, Jinja

- ***AI/ML toolkits & Algorithms***: CNN
- ***AI/ML Libraries***: OpenCV-python
- ***DB***: MySQL
- ***Version control***: GITLAB




Hardware Requirements

- Windows 10
- A smartphone (Android 7.0)
- Intel® Core™ i7-8250U Processor
- 8 GB RAM

4. Project Requirements

- **4.1 Project Plan**

Research and Planning:

The first step in the project plan is to conduct thorough research and planning. This involves identifying the key goals and objectives of the project, determining the scope of the project, and researching the best approaches and technologies for sign language recognition. This stage also involves identifying the necessary resources and personnel for the project.

Data Collection: Once the research and planning stage is completed, the next step is to collect a diverse dataset of videos containing the sign language actions to be recognized. This dataset should include a variety of sign languages, dialects, and variations to ensure that the machine learning model is trained on a comprehensive set of data.

Data Preparation: The collected data must then be prepared for training the machine learning model. This involves labeling each video with the correct action it contains, segmenting the videos into shorter clips, and extracting visual features from the frames in each clip. Accurate labeling and feature extraction are crucial for the effectiveness of the model.

Model Training: The machine learning model is then trained using the labeled and prepared data. This typically involves using deep learning algorithms, such as CNNs or RNNs, to learn patterns in the visual features extracted from the videos. The model is trained to classify each clip into one of the sign language actions it represents.

Model Testing and Optimization: Once the model has been trained, it is tested on a separate set of videos to evaluate its accuracy and performance. The model is then optimized to improve its accuracy and effectiveness in recognizing sign language actions. This involves adjusting the parameters of the model, such as the learning rate or the size of the layers, to improve its performance.

Integration with Ecommerce Site: Once the model has been trained and optimized, it can be integrated into the ecommerce site. This involves developing the necessary software and interfaces to allow customers to use sign language to communicate with the site.

User Testing and Feedback: After the integration of the sign language recognition feature, user testing and feedback are essential to evaluate its effectiveness and user experience. Feedback can be used to further optimize the system and ensure that it meets the needs of customers.

Maintenance and Improvement: The final step in the project plan is to maintain and improve the sign language recognition feature. This involves regularly monitoring and updating the system to ensure that it continues to perform effectively and efficiently. Additionally, updates and improvements can be made based on user feedback and advances in technology.

4.2 Project Management

- Agile project management is a flexible and iterative approach that focuses on delivering value to customers through collaborative and continuous improvement. Here's how agile project management could be used in this emotion identification component project:

Sprint Planning

- Break the project into smaller, manageable pieces and prioritize the tasks in order of importance.
- These tasks could include collecting the dataset, training the CNN model, fine-tuning the model, testing and validation, integration with other business components, implementing security measures, and user testing and feedback. Each of these tasks should be planned for in each sprint.

Daily Stand-Ups

- Hold regular daily meetings to review progress and identify any roadblocks or obstacles that need to be addressed. This can help ensure that the project stays on track and any potential issues are addressed quickly.

Sprint Reviews

- Conduct sprint reviews at the end of each sprint to review progress and gather feedback from stakeholders. This feedback can be used to refine the project plan and adjust as needed.

Continuous Improvement

- Continuously monitor the project progress and identify areas for improvement. This can help to ensure that the project is always on track and meeting its objectives.

Collaborative Approach

- Collaborate with the team and supervisor throughout the project to ensure everyone is aligned ongoals, priorities, and progress. This can help to ensure that everyone is on the same page and working towards a common objectives

- **Functional Requirements:**

Recognition of a wide variety of sign language actions: The system must be able to recognize a broad range of sign language actions to enable users to communicate effectively with the ecommerce site.

Accuracy: The system must have high accuracy in recognizing sign language actions to avoid miscommunication or frustration for users.

Speed: The system must operate quickly, with minimal lag time between sign language input and the corresponding action on the ecommerce site.

Integration with the ecommerce site: The system must be seamlessly integrated into the ecommerce site to allow users to navigate and complete tasks using sign language.

Adaptability to user preferences: The system must be adaptable to different sign language dialects and variations to accommodate a diverse range of users.

User feedback and improvement: The system must allow for user feedback and incorporate improvements based on user input to enhance its effectiveness and user experience.

- **Non-Functional Requirements:**

Accessibility: The system must be accessible to all users, regardless of their communication preferences or abilities.

Security: The system must be secure to protect users' personal and financial information.

Reliability: The system must operate reliably, with minimal downtime or errors, to avoidfrustrating or inconveniencing users.

Scalability: The system must be scalable to accommodate a growing number of users and anincreasing amount of sign language input.

User privacy: The system must protect users' privacy and data, complying with relevant privacylaws and regulations.

Usability: The system must be intuitive and easy to use, with clear instructions and userinterfaces to enable users to navigate and use the system effectively.

Commercialization Plan

It is an e-commerce platform that offers a cutting-edge shopping experience for the Deaf and Hard of Hearing (DHH) community and provides an AI-powered voice assistant for all users. Our platform includes sign language interpretation, AI-driven product recommendations, and a seamless shopping experience To make online shopping more accessible and enjoyable for everyone, regardless of their communication preferences or abilities, by leveraging advanced AI technology. DHH individuals and their families. Tech-savvy shoppers looking for a convenient, voice-assisted shopping experience.Online advertising through social media and Google Ads Partner with DHH organizations and influencers for promotion.Targeted online marketing campaigns. Engage with potential users through social media and email marketing

5.3 Future work

For future developments, it can be enhanced by developing this system containing more hand gestures. And User testing

Conclusion

In conclusion, the integration of sign language recognition into digital systems is a promising approach to enhancing accessibility for differently abled individuals. As the digital age continues to evolve, it is crucial to ensure that technology is accessible to everyone, regardless of their individual needs and requirements. By incorporating sign language recognition into e-commerce, we can provide a more inclusive and accessible online shopping experience for individuals who use sign language as their primary means of communication. Our proposed workflow can improve the independence and efficiency of differently abled individuals and promote equal access to digital systems. While current research has mainly focused on improving accessibility for individuals with specific disabilities, our approach aims to provide a more inclusive and accessible online shopping experience for all. By promoting inclusivity and accessibility in e-commerce, we can foster a more diverse and equitable society that values and supports the participation of individuals with disabilities. It is also essential to have a team of experienced data scientists and developers with expertise in machine learning, computer vision, and web development to design, implement, and maintain the sign language recognition system and its integration into the ecommerce site.

Used Hand Gestures

- **Previous**

- **Next**

- **Cart**

- **ChatBot**

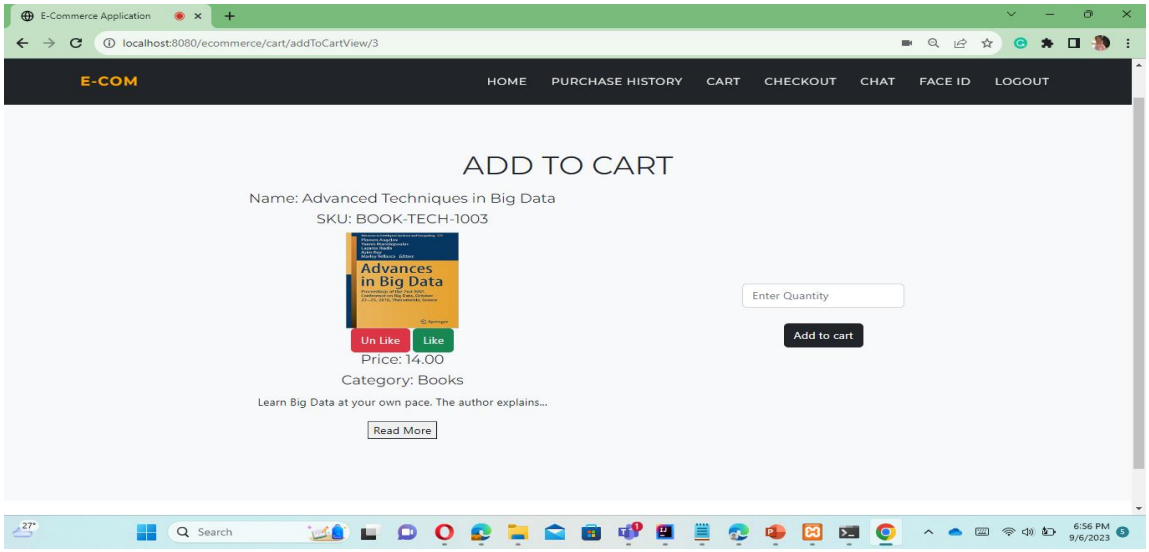
- **Checkout**

- **Scroll up & Down**

Sample Code for hand gesture training

```python
@app.route('/number-predictor-api', methods=["POST"])
def number_predictor_api():
    try:
        # Load and Save the Image
        image = request.files["image"]
        # Getting file name of the image using werkzeug library
        filename = secure_filename(image.filename)

        if image and allowed_file(image.filename):
            # Saving the image in temp directory
            image.save(os.path.join(app.config["UPLOAD_FOLDER"] + "/temp/number-predictor", filename))

            frame_path = "images/temp/number-predictor/image.PNG"
            landmarks = get_hand_landmarks(frame_path)
            if landmarks:
                # Create a list to store landmark dictionaries
                landmark_data = []

                # Append landmark coordinates to the list
                for landmark in landmarks:
                    landmark_data.append(landmark[0])
                    landmark_data.append(landmark[1])

                # Convert the list to a DataFrame
                predictions = loaded_model_number_predictor.predict([landmark_data])
                os.remove(frame_path)
                return str(predictions[0])
            else:
                os.remove(frame_path)
                return "No number found"
        else:
            return "Images are not of the allowed extensions!"
```

Expected Output



REFERENCES

[1] Ge, L., Li, Y., Zhang, J., & Yang, J. (2012). Real-time sign language recognition using a Kinect sensor. International Journal of Distributed Sensor Networks, 8(3), 139383.

[2] Fan, J., Zhou, K., Lv, C., Liu, X., & Zhou, Z. (2017). Sign language recognition using recurrent neural networks with long short-term memory. IEEE Access, 5, 12883-12890.

[3] Li, J., Lin, W., Li, X., & Yang, R. (2018). Sign language recognition based on dynamic features and deep learning. IEEE Access, 6, 33061-33071.

[4] Haque, R., Islam, S., Khan, M. R. I., & Chowdhury, M. S. (2019). Sign language recognition using 3D convolutional neural networks. 2019 International Conference on Robotics and Automation (ICRA), 751-757.

[5] Li, J., Zhang, J., Yu, W., Yang, J., & Gao, X. (2021). A wearable system for real-time sign language recognition using surface electromyography. IEEE Access, 9, 116568-116579.

[6] Al-Tamimi, J., & Azmi, S. (2019). Sign language recognition using image processing and machine learning: a systematic review. Journal of Intelligent Systems, 28(1), 23-36.

[7] Nguyen, L. M., Nguyen, H. T., Nguyen, L. T., Nguyen, T. T., & Vo, N. H. (2020). Sign language recognition: A comprehensive review. IEEE Access, 8, 2434-2462.

[8] Han, Y., He, L., Ji, Y., & Sun, Y. (2020). Sign language recognition based on deep learning: A survey. IEEE Transactions on Neural Networks and Learning Systems, 32(4), 1247-1262.

[9] Wang, W., & Xu, Y. (2020). Sign language recognition: A survey. arXiv preprint arXiv:2002.12439.

[10] Pratama, G., Tan, J. K., & Er, M. J. (2021). A review of sign language recognition using deep learning. IEEE Access, 9, 31648-31666.

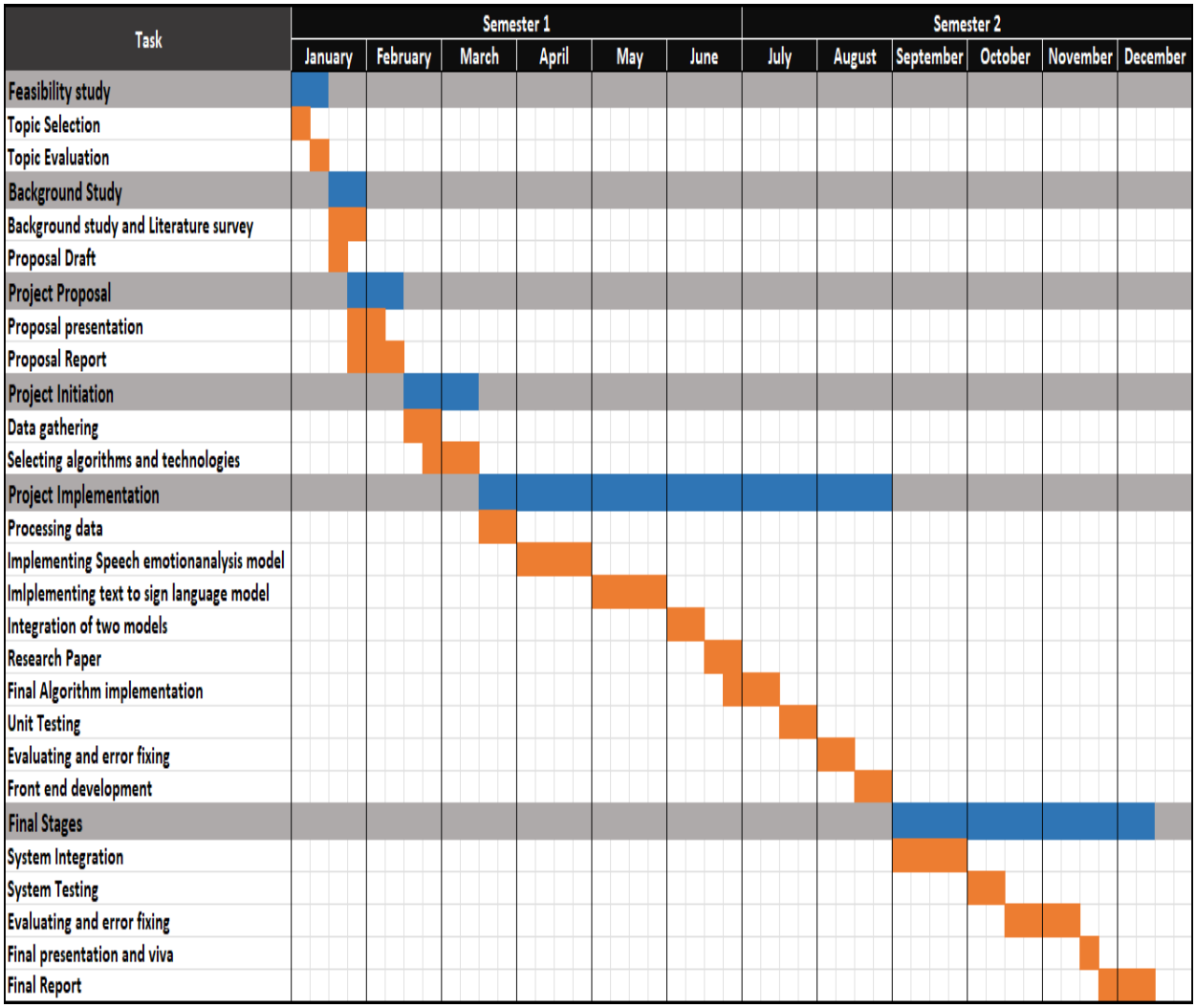
[11] J. R. Johnivan, “Best Agile Development Methodology & Principles for 2022,” Project-Management.com, Nov. 19, 2022. https://project-management.com/agile-software-development-methodologies

[12] Kim, K., Kim, K., Hong, S., & Kim, H. (2018). Wearable haptic device for screen navigation using sign language. IEEE Access, 6, 6381-6391.

[13] Naim, U. I., Islam, M. R., & Shatabda, S. (2019). Deep learning-based sign language recognition and translation for screen navigation. In 2019 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.

[14] Wu, Y., Sun, J., Chen, J., Zhou, M., & Xu, H. (2020). A wearable device-based screen navigation system for deaf-mute people using sign language. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(5), 1166-1177.

Gannt Chart


Task
Semester 1
Semester 2
January
February
March
April
May
June
July
August
September
October
November
December
Feasibility study
Topic Selection
Topic Evaluation
Background Study
Background study and Literature survey
Proposal Draft
Project Proposal
Proposal presentation
Proposal Report
Project Initiation
Data gathering
Selecting algorithms and technologies
Project Implementation
Processing data
Implementing Speech emotionanalysis model
Imlplementing text to sign language model
Integration of two models
Research Paper
Final Algorithm implementation
Unit Testing
Evaluating and error fixing
Front end development
Final Stages
System Integration
System Testing
Evaluating and error fixing
Final presentation and viva
Final Report

Work Bench Chart