HAWK.AI: Revolutionizing Dining Experiences with Touchless Gesture Recognition Technology

Ashlin Joel, Navamani T M, Archita Sharma, Vansh Harkut, Mansi Saxena and Arnav Goenka

***Abstract*—In the current digital age, the need for contactless interaction has become increasingly important. To address this need, our research provides an innovative solution that uses advanced technology to prevent the spread of germs in busy areas like ATMs and various places that engage touch screens, especially aftermath of the COVID-19 pandemic. Our work introduced a touchless digital menu system using advanced hand gesture recognition technology to reduce the risk of spreading germs and serve as a boon to help hearing-impaired & speech-impaired people. This research focused on a smart algorithm of human-computer interaction and gesture recognition that could recognize eleven different gestures using finger counting. By smoothly combining picture taking, pre-processing, finding fingers, and knowing gestures, our system shows accuracy at identifying gestures of 1 to 5 fingers. This adaptable solution addresses immediate pandemic concerns and offers a forward-looking approach to future challenges. Also, our touchless system allows inclusive interaction for people with hearing or speech problems, breaking down communication barriers. This research work contributes to ongoing efforts to decrease viral spreading and offers a solution using advanced technology.**

***Index terms*—Touchless digital menu system, Hand motion recognition technology, Gesture recognition, Human-computer interaction**

I. INTRODUCTION

In the ever-changing realm of computer hardware and software, we live in a time that promises more engaging user experiences. Notably, physical gesture language emerges as a powerful and effective communication tool frequently neglected in conventional human-computer interaction. There was a notable surge in the need for accurate and effective recognition of hand gesture recognition, the process of interpreting human gestures via mathematical algorithms to interact with digital devices. Human activities in smart environments were being recognized by sensors that detect depth, analyzing both spatial and temporal activity patterns, and applying a statistical model called a hidden Markov model to understand and predict these activities [1]. It involves capturing, analyzing, and understanding gestures made by body parts or devices like motion sensors. Key applications include Gesture-based gaming control and gesture-controlled car driving. These results highlighted Human-Computer Interaction (HCI) where gestures replace traditional input methods, enabling intuitive control in gaming, virtual reality, and smart homes. In healthcare, gesture recognition aids in rehabilitation and surgery by tracking movements. Automotive systems utilize gesture recognition for hands-free control of infotainment and navigation. Security systems employ it for biometric authentication.

The motivation behind choosing this work stemmed from the pressing need to address the challenges associated with touchscreen-based menus, as highlighted by respondents in our analysis. The decision to pursue this work was driven by the significant interest and positive feedback received from respondents regarding the potential of a gesture-based food ordering system. With 21.1% strongly agreeing and 34.2% agreeing that such a system could address the challenges associated with traditional and touchscreen-based ordering methods, it became evident that there was a demand for innovative solutions. These responses underscored the motivation to explore and develop a gesture-based system, highlighting its potential to enhance user experience and alleviate the shortcomings of existing ordering methods, as evidenced by the improved human activity recognition achieved through smart windowing and spatio-temporal feature analysis [2].

The process begins with the challenging process of distinguishing and identifying hands in photos, which is made more difficult by the different lighting conditions, scales, orientations, and positions of the hands. Digital image processing is at the forefront of our investigation, a dynamic force driving the software sector. The promise of introducing machines that mimicked human visual processes lay in their potential applicability across a range of domains. Different Approaches for Human Activity Recognition are constantly being made across various use cases across various domains from video surveillance to health monitoring [3][6]. Gesture recognition systems offer a more seamless and user-friendly experience in Human activity recognition in RGBD videos by dynamic images offer insights into advanced technological approaches [4]. Hand gesture recognition is one of the key biometric technologies that are becoming more and more popular in the private and law enforcement sectors due to the growing demand for security applications.

In recent years, the research has identified several challenges in hand gesture recognition, including limited gesture vocabulary, varying lighting conditions, occlusion, real-time processing constraints, user-dependent variations, and integration with existing interfaces. Ferrari et al. [7] proposed personalized classification models for Human Activity Recognition using accelerometer signals from smartphones. Similarly, some of the other previous studies addressed lighting variability with adaptive algorithms and image pre-processing techniques. We aim to enhance gesture recognition by leveraging advanced algorithms and machine learning techniques to expand gesture vocabulary, adapt to varying conditions, and accommodate user-specific variations. By integrating depth sensing technologies and optimizing real-time processing to provide robust and seamless gesture recognition experiences across diverse environments with high user populations. Through seamless integration with existing interfaces, our work endeavors to enhance user interaction experiences and address the identified challenges in hand gesture recognition.

The existing research works often faced issues and drawbacks in hand gesture recognition methodologies. Many approaches struggled with accuracy, especially in complex environments with varying lighting conditions and backgrounds. Some of the existing works relied on traditional computer vision techniques, which often lacked robustness and failed to generalize well across diverse hand gestures.

Our work addresses these limitations by employing a hybrid approach that combines machine learning, and artificial intelligence with computer vision techniques. We utilize the MediaPipe library for hand tracking and landmark detection, allowing for precise identification of hand gestures. This system works well and offers a well-designed and documented base that can be adjusted to various input sizes and resolutions Unlike some previous methods that struggled with real-time performance, our system maintained high accuracy while ensuring efficient processing, crucial for interactive applications like restaurant food ordering. By integrating advanced algorithms and leveraging deep learning models, the proposed system can accurately interpret a wide range of hand gestures, even in challenging environments. Moreover, the said work emphasizes user experience by offering intuitive and seamless interaction, enhancing the overall dining experience. Through rigorous testing and validation, this work demonstrates the effectiveness and reliability of our approach, paving the way for more practical and user-friendly gesture recognition systems. Dedicated to simplicity, this system's brains were contained in a single workstation or computer, negating the need for specialized hardware.

The subsequent sections of the research paper are as follows: Section II provides the background by evaluating the previous studies and identifying the shortcomings that will be addressed by the proposed approach. Section III involves survey designing, data collection, and analysis which are essential for the system development. Section IV shows how computer vision is combined with web development to have hand tracking, gesture recognition, and website integration. Section V is dedicated to the performance of the gesture recognition algorithm and the implications for user experience improvement. Section VI brings together the main findings, discusses the prospects for further research, and proposes how gesture recognition technology can change how we dine in restaurants.

II. LITERATURE REVIEW

The advancement in gesture recognition and Human Activity Recognition (HAR) has garnered significant attention in recent years due to its multifaceted applications and implications across various domains. Mohamed et al. [8] highlighted the ongoing need for further advancements in continuous gesture recognition to develop practical vision-based systems capable of discerning subtle and dynamic movements accurately. Oudah et al. [9] categorize hand gestures into dynamic and static types, emphasizing their varied applications in facilitating Human-Machine Interaction (HMI) and enhancing user experience in interactive systems. Guo et al. [10] underscored the pivotal role of Hand Gesture Recognition (HGR) in facilitating intuitive communication, citing its high efficiency in conveying information with minimal latency.

Al-Hammadi et al. [11] emphasized the growing significance of automatic hand gesture recognition, particularly in addressing the needs of the hearing-impaired population and advancing touchless control technologies in various devices and applications. Parvathy et al. [12] presented a robust HGR system that effectively tackled complex background issues, enhancing overall recognition performance by leveraging sophisticated machine learning algorithms and feature extraction techniques. Mujahid et al. [13] proposed a model that was capable of achieving high accuracy even in challenging environments and low-resolution scenarios, highlighting the robustness and adaptability of deep learning-based approaches in real-world settings.

Zhang et al. [14] delved into the naturalness of hand gestures in human-robot interaction scenarios, emphasizing the importance of vision-based dynamic gesture recognition methods in enhancing user experience and enabling seamless communication between humans and machines. Sharma et al. [15] stressed the growing importance of image identification in modern problem-solving systems, underscoring its relevance in various applications, including HGR systems deployed in smart environments. Xu et al. [16] explored the limitations of traditional feature engineering methods in handling the complex waveform data characteristic of HAR systems, advocating for the adoption of deep learning models to overcome these challenges and improve recognition accuracy and efficiency.

Bevilacqua et al. [17] defined HAR as the automatic identification of physical activities performed by humans, highlighting its significance in healthcare monitoring, sports analysis, and other domains. Wan et al. [18] addressed the limitations of traditional methods in identifying complex and real-time human activities, advocating for the adoption of deep learning models for improved performance and scalability. Ramanujam et al. [19] provided a comprehensive review of deep learning techniques for HAR using smartphone and wearable sensors, emphasizing the need for further research to address existing challenges, such as data heterogeneity and model interpretability, and enhance system efficacy in real-world applications.

Based on the wide variety of findings brought out by the extensive spotlight review process conducted, the goal of the work was to address the flaws identified in past systems. It aims to provide a comprehensive solution that allows customers to place restaurant food orders in one place, thus surpassing the numerous setbacks of the older versions. In this case, a greater range of capabilities beyond the standard menu navigation functionalities needs to be considered, such as real-time status updates of orders, allergy notifications, and efficient coordination with kitchen workers. By supporting the implementation of novel technologies and taking advantage of the errors of earlier systems, the site will be privileged to achieve its aim, which is to provide a dining experience that is augmented and streamlined for its users.

III. RESEARCH METHIDOLOGY

A questionnaire was an optimal choice for our research methodology compared to other methods due to various advantages. Due to its simplicity in reaching a diverse group of people, it allowed us to gather data from a larger and more diverse sample size. Since the names of participants were not asked it provided participants with anonymity and freedom to express their true opinions, preventing any bias that may arise due to peer pressure. The questions were short, crisp, and to the point and the questionnaire was not time-consuming to fill which proved a better alternative than other options available which could have been more time-consuming and might not fit the attention span of people. Moreover, they were cheaper, requiring fewer resources as compared to methods like interviews or experiments. The purpose of the questions was to guarantee uniformity in the gathering of data, minimize biases, and improve the dependability of the results. Additionally, questionnaires made it easier to gather and analyze data efficiently. Kaur et al. [20] proposed an AI trainer model utilizing Personal Status Evaluation and OpenCV for real-time human activity recognition, facilitating fitness training regardless of age or health status, which would be a cost-effective alternative to wearable and portable devices, because of its adaptability and flexibility, changes could be made by answering new research topics or by considering user feedback, ensuring that the study remains current and sensitive to the changing user demands. The questionnaire proved to be a useful and effective method for obtaining important information for our research on creating an inclusive gesture-based food ordering system.

A Google Form containing 10 questions of interest was devised to comprehend user expectations and requirements and the list of these 10 questions can be seen in Table 1. The questions were crafted to acquire insights into users' dining habits, preferences, the challenges they had encountered with traditional food ordering systems, and their perspectives regarding integrating a gesture-based system into food ordering.

3.1 ANALYSIS

A comprehensive survey was conducted, recording 40 responses predominantly from students of VIT Vellore University, to gain insights into dining habits and preferences. The primary objective of the survey was to identify the challenges associated with traditional ordering methods namely physical menus, QR-based menus, and touchscreen-based food ordering systems, and to evaluate the feasibility of gesture-based solutions. The survey commenced with a general question designed to determine the frequency with which individuals dine out, yielding the following results:

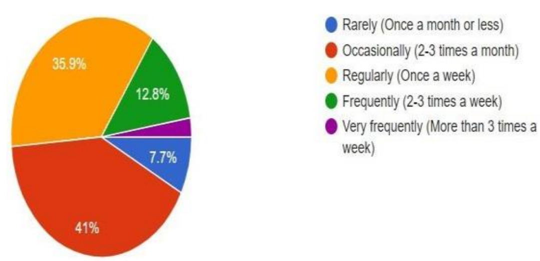


Fig. 1. Frequency of dining out by people

Thus, based on Figure 1, we can infer that most individuals frequently eat out, which suggests that our product has a wide market.

TABLE I

QUESTIONNAIRE USED DURING THE SURVEY FOR ANALYSIS OF THE TRADITIONAL RESTAURANT ORDERING METHODS.

QUESTIONS	RESPONSE
How often do you go out to dine in restaurants?	• Rarely (Once a month or less) • Occasionally (2-3 times a month) • Regularly (Once a week) • Frequently (2-3 times a week) • Very frequently (More than 3 times a week)
What factors do you consider important for a good dining experience?	• Quality of food • Ambiance/atmosphere • Service Quality • Price/value for money • Convenience of food ordering • Other: (Manual Response)
What are some challenges you have experienced with the traditional food ordering system (using physical menus) in restaurants?	• Long waiting times • Difficulty in finding desired items • Miscommunication with staff • Limited visibility of menu options • Other: (Manual Response)
Are there any concerns or drawbacks you associate with QR code-based menus for food ordering?	• Difficulty in scanning QR codes • Concerns about data privacy • Limited accessibility for some users • Dependence on Internet connectivity • Other: (Manual Response)
What issues do you face with touchscreen-based food ordering systems?	• Hygiene concerns as the same touchscreen is being used by many users • Technical glitches, lags, or malfunctions • Difficulty in navigating the interface • Sensitivity to touch leading to unintentional orders • Other: (Manual Response)
How satisfied are you with the current technological advancements in food ordering systems? (1 being very dissatisfied, 5 being very satisfied)	The option to rate was given on a range from 1 to 5.
Do you believe a gesture-based food ordering system (wherein you can navigate through the menu through your gestures) could address the problems associated with traditional and touchscreen-based ordering methods? (1 being strongly disagree, 5 being strongly agree)	The option to rate was given on a range from 1 to 5.
What features do you consider essential to be displayed when ordering food through a digital menu (touchscreen-based menu/ QR based menu or gesture-based menu)?	• Images of dishes • Description/ingredients of dishes • Trending/Chef's special dish highlights • Allergy warnings/Spicy warnings • Vegetarian/Non-vegetarian indicators • Estimated preparation time • Add your preferences(like without onion/garlic etc.) • Other: (Manual Response)
How likely are you to adopt a gesture-based food ordering system if it were available in restaurants? (1 being very unlikely, 5 being very likely)	The option to rate was given on a range from 1 to 5.
What potential problems or challenges do you foresee with the implementation of a gesture-based food ordering system in restaurants? (Please specify)	• Difficulty in accurately interpreting gestures • Limited accessibility for individuals with certain disabilities • Concerns about privacy and data security • Technical glitches or malfunctions • High implementation costs • Resistance to change from traditional ordering methods • Other:

Findings revealed prevalent issues with physical menus, including extended waiting times (68%), difficulty in finding desired items (47%), and miscommunication with staff (42%). Thus, creating systems that reduce employee involvement and provide intuitive user interfaces for easy navigation and speedy order processing became necessary. The main issues with QR-based menus were dependence on Internet connectivity (63%), data privacy (42%), and limited accessibility (40%). Ensuring robust Internet connectivity and implementing stringent data privacy measures emerged as critical priorities.

Regarding touchscreen-based menus, respondents equally highlighted technical glitches, lags, and malfunctions (58%), along with challenges in navigating interfaces (58%). Additionally, concerns about unintentional touches due to screen sensitivity (40%) and hygiene issues (34%) were noted. The proposed gesture-based system held promise in mitigating touch-related concerns and enhancing user experience through fluid and responsive interactions.

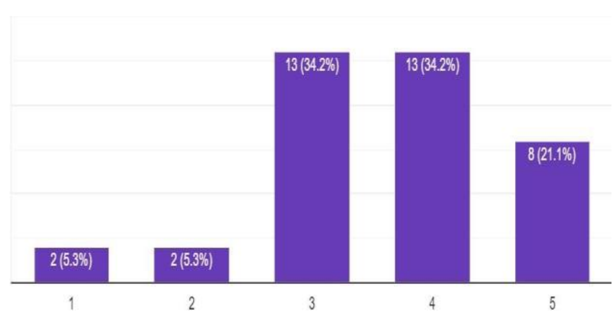


Fig. 2. People's opinion on the efficiency of a gesture-based system in solving problems present in old traditional food ordering system

Next, we gathered people's opinions on this gesture-based system's ability to mitigate all the problems they faced with traditional food ordering methods. (here 1 meant strongly disagreed and 5 meant strongly agreed).

As evident in Figure 2, participants expressed a favorable inclination towards gesture-based technology as a solution to existing challenges, although some apprehensions were evident. Following that, we inquired about the expectations and requirements of individuals, specifically regarding the features they anticipated in a digital food ordering system. The anticipated features can be seen in Figure 3.

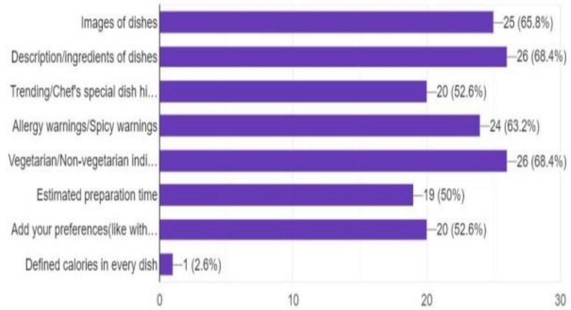


Fig. 3. Essential features to be incorporated into the system.

Additionally, survey participants showed they are open to using gesture-based ordering systems, though some hesitated due to the novelty of the technology as evident from Figure 4 (1 represents strongly disagree while 5 represents strongly agree)

Whereas Figure 5 showcases concerns identified with gesture-based systems, issues primarily revolved around the accuracy of gesture recognition, highlighting its critical role as a unique selling proposition. Ensuring bug-free operation, smooth functionality, and robust security measures were key to the success of the proposed system.

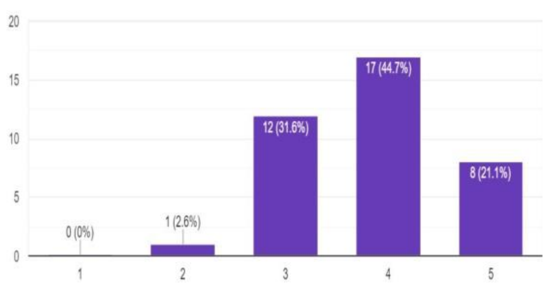


Fig. 4. People's readiness to adopt new technology

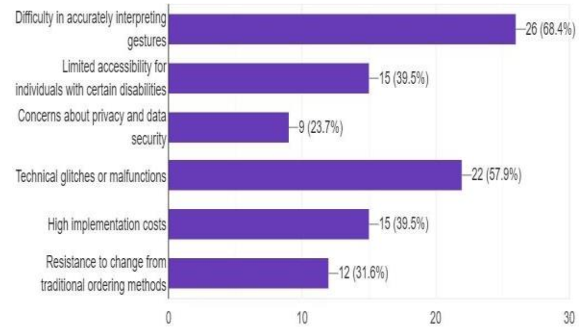


Fig. 5. Limitations with the gesture-based system

In conclusion, the survey findings underscored the importance of prioritizing gesture recognition technology and system reliability in developing innovative food-ordering solutions.

IV. ALGORITHM DEVELOPMENT

The Algorithm Development section explains the process of making and improving the algorithm that is critical to the operation of a gesture-based food ordering system. This important step involves the integration of advanced computer vision techniques and web development models to enable real-time hand tracking and gesture recognition. Through careful testing and incremental development, the algorithms presented here demonstrate their effectiveness in accurately and efficiently interpreting user gestures. This section explores the technical variations of algorithmic design and highlights the fusion of innovative methods to achieve seamless interaction between users and the digital interface. In addition, it highlights the transformative potential of these algorithms to change traditional food paradigms, highlighting their role in improving the user experience and at the frontier of human-computer interaction.

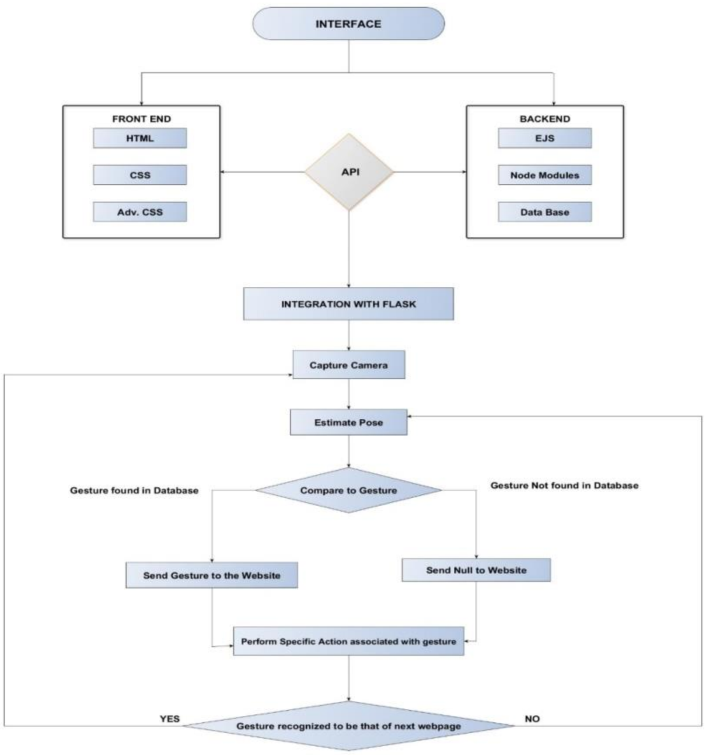


Fig. 6. Block diagram for the entire workflow model of the proposed system.

4.1 Software Tool Description

A crucial part of this research work was the Algorithm-building phase, which describes the painstaking integration of state-of-the-art computer vision techniques with web-building methodologies. This section explains how to employ gesture recognition technology to improve user experience on a website that serves digital restaurant menus.

In this regard, it afforded us with means to scale up and gain a broader reach that let us investigate the topic of users' tastes and actions from a much wider sample, thus receiving almost full and comprehensive information. It provided a means to scale up and gain a broader reach, allowing for the investigation of users' tastes and actions from a much wider sample. By integrating the front end of the webpage made of HTML, CSS, and JavaScript elements with Flask, a Python web framework. The algorithm makes it easier to navigate, choose dishes, and control carts—all of which improve the overall ordering and dining experience.

4.2 Gesture Recognition Algorithm:

4.2.1 Initialization and Setup:

The algorithm initializes the hand detection model using the MediaPipe library, configuring parameters such as model complexity, maximum number of hands to detect, and detection confidence thresholds. It also sets up the video capture device to access the webcam feed for real-time input.

4.2.2 Palm Detection and Hand Tracking:

The algorithm continuously captures frames from the webcam feed and processes them to detect hands using the Mediapipe Hand module as shown in Figure 7. For each frame, it converts the image to RGB format and passes it through the hand detection model to identify hand regions.

Detected hand regions are tracked across frames to maintain continuity and stability in hand localization. The suggested solution is based on hand tracking, which allows for the real-time identification and monitoring of the user's dexterous hand gestures using video stream analysis. The application employed powerful computer vision libraries, such as Mediapipe, to precisely recognize essential hand landmarks, a total of 21 of which have been pointed out clearly in Figure **7**. This enabled accurate localization and tracking during the contact process. This allows hand motions to be easily understood and turned into commands for use in the digital menu interface. The HandDetector class in the Python code creates a hand-tracking object by utilizing the Mediapipe module. Every webcam image is processed by this class, which then tracks and identifies important hand landmarks like joints and fingertips. The findings underscore the significance of Joo, Simon, and Sheikh's research in advancing multi-modal human pose estimation techniques, paving the way for more accurate and comprehensive tracking systems.

4.2.3 Hand Landmark Extraction:

Once hands are detected, the algorithm extracts landmarks representing key points on each hand, such as fingertips, knuckles, and wrist. It iterated through the detected hand regions, retrieving the 3D coordinates of each landmark relative to the image dimensions. Landmark positions are stored in a list data structure for further processing.

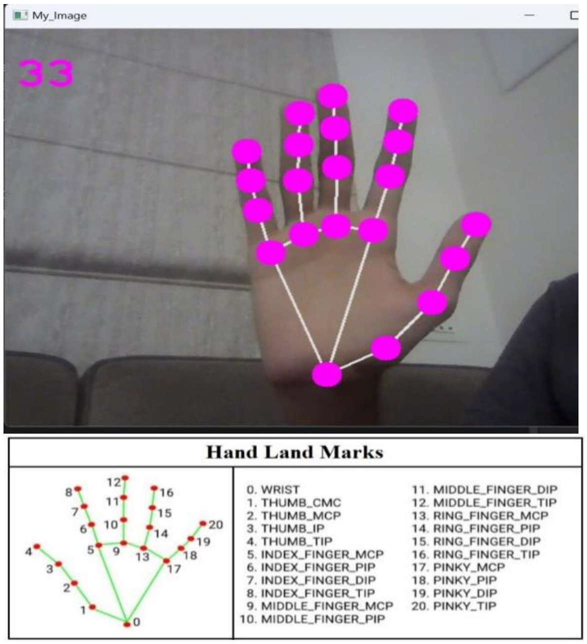


Fig. 7. Showcasing the 21 hand landmarks

4.2.4 Gesture Recognition Logic:

Here, we implemented Gesture Recognition logic to recognize specific hand gestures based on the positions of landmarks. Gesture recognition involves identifying patterns in landmark positions indicative of predefined gestures, such as "Cheers," "Rock," "Love," etc. which can be seen in Figure 8. Therefore, one of the essential components of the algorithm is gesture recognition, which converts hand gestures into meaningful actions on the digital menu screen. By carefully examining hand landmarks and spatial arrangements, the system recognized preset motions, enabling users to choose food, traverse menu options, and interact with the website with ease. Sophisticated algorithms are used in this process to identify motions with a high degree of precision and consistency by analyzing hand landmarks. Our approach prioritized intuitive and user-friendly gesture recognition for navigating the digital menu interface.

4.2.5 Overlaying Visual Feedback:

Upon recognizing a gesture, the algorithm overlays corresponding images and text on the video feed to provide visual feedback, this has been depicted above in Figure 8. It dynamically updates the display with the recognized gesture and associated feedback, enhancing user interaction and experience. Visual feedback was superimposed on the video frame using OpenCV drawing functions, ensuring real-time responsiveness.

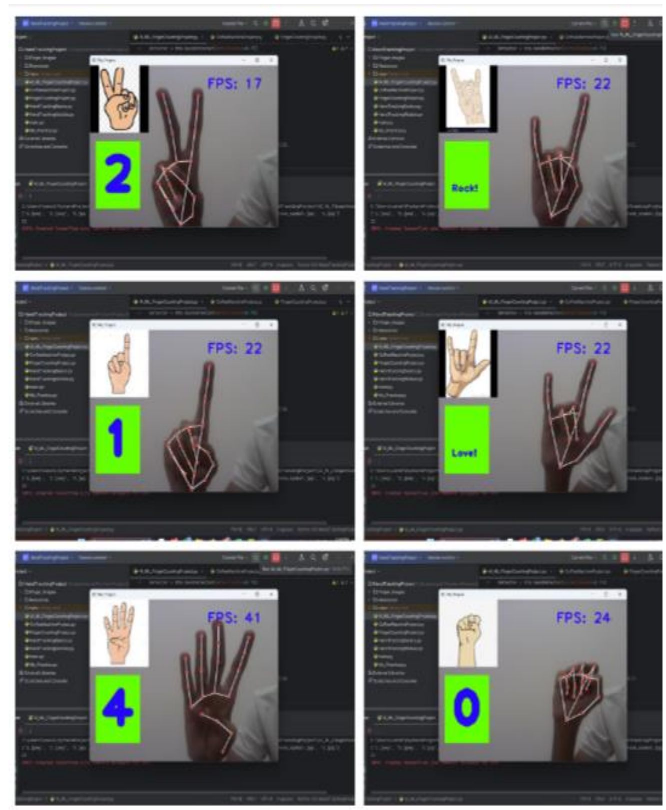


Fig. 8. Showcasing gesture recognition based on the Hand landmarks

4.2.6 Flask Backend Integration:

The gesture recognition algorithm is seamlessly integrated with Flask, a lightweight and efficient Python web framework, to create a robust backend server. Flask routes are defined to handle different Hyper Text Transfer Protocol (HTTP) requests originating from the frontend User Interface. For instance, routes are established to capture webcam input, process gesture recognition requests, and serve the processed results back to the front end. The backend server leverages Flask's simplicity and flexibility to manage the flow of data between the front end and the gesture recognition algorithm. It efficiently processes incoming requests, executes the gesture recognition logic, and returns relevant information to the front end promptly. Utilizing Flask's modular structure, the backend code is organized into logical components, ensuring maintainability and scalability as the work evolves. Each route corresponds to a specific functionality, facilitating easy debugging, testing, and future enhancements.

4.3 Integration with Website

4.3.1 Design

The website was rigorously designed to match the changing needs of the restaurant industry, with an emphasis on interactivity. Its distinctive use of Hyper Text Markup Language (HTML), Cascading Style Sheets (CSS), and Advanced CSS codes gave websites a modern appearance. JavaScript, the fundamental programming language, powers functionality. Embedded JavaScript (EJS) ensures that the front and back ends work together flawlessly on a variety of websites.

The application's backend infrastructure uses Postgres ŞQL, a robust database system. Data management and retrieval operations become more efficient as a result. Furthermore, the integration of the Express and Node.js frameworks as the Application Programming Interface (API) medium allows for easy communication between the website's front and back ends. A detailed block diagram representing this can be seen in Figure 6**.** Only with this tactical combination of technologies can the diverse demands of the restaurant business be handled with optimal performance and scalability.

4.3.2 Implementation and Evaluation

To efficiently handle the ordering process, our application makes use of two different databases. Included in the main database, called "menu," were vital details such as the dish's name, price, and unique identification. An image depicting the name and price of the dish upon being added to the cart can be seen in Figure 10. However, we do have a temporary database called "placedorders," which is intended to hold the meals that customers have chosen for their orders. Notably, to easily handle following orders, this temporary database is cleared at the end of each order cycle.

Let's look at an example of how a consumer would use our application to make an order at a restaurant to show the user experience. Initially, the camera interface records the user's hand gestures as they explore the menu to reach the non-vegetarian food selections. Every dish has a unique gesture connected to it on the screen, making selecting simple. After that, the customer can select the dishes they want from a variety of categories, including drinks, appetizers, main courses, and desserts.

Using a specified gesture, the user can easily review the products they had selected in the cart during the selection process. The items that have been selected and added to the cart can be seen in Figure **10**. Before completing their order, individuals can choose to eliminate any dishes they do not want. The above description showcasing the complete workflow has been represented in Figure 9.

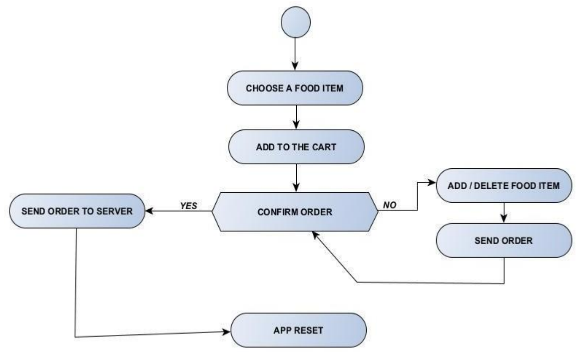


Fig. 9. The internal functionality of the website after the implementation of gesture recognition.

Each option causes a form to be submitted in the background, with a hidden input associated with the dish's key ID in the main menu database. After submission, the backend parsed the page and used the key ID to extract the related dish details from the main menu database. This data is then kept in the temporary database, which makes it easier to read and write across the two databases. The contents of the temporary database are retrieved and shown when the user wants to see their cart. This allows the user to review their choices and make any required corrections before moving forward.

The temporary database's contents are erased upon order confirmation, effectively managing the current order without affecting the main database. Throughout the ordering process, this procedure preserves data integrity and efficiency while guaranteeing a smooth and efficient experience for users.

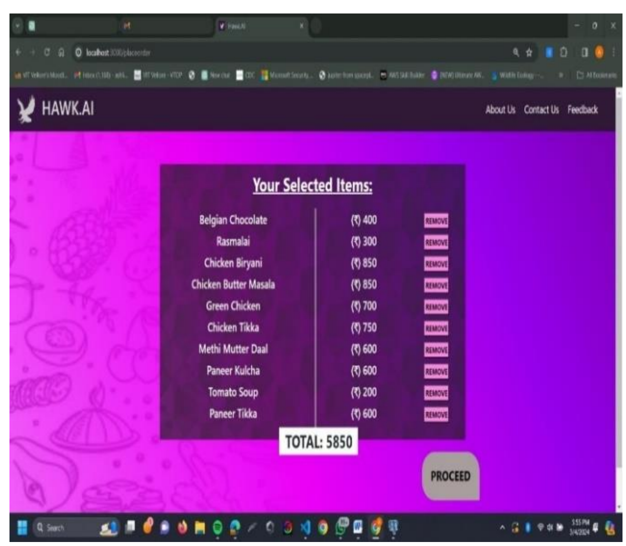


Fig. 10. A diagram of items that have been added to the cart on the website.

4.3.3 Seamless Backend-Frontend Interplay

Flask's integration with the HTML, CSS, and JavaScript website creates a smooth and easy-to-use user experience by promoting a balanced interaction between frontend and backend processing. As the foundation of the integration, Flask—which is well-known for its ease of use and adaptability—manages the reception of gesture data from the client-side application and initiates the necessary actions on the website.

4.3.4 Flask Endpoint for Gesture Handling

The primary component of this connection is a Flask endpoint that manages incoming gesture data. After the client-side application sends gesture data in JavaScript Object Notation (JSON) format, Flask uses the advanced hand-tracking and gesture-detection algorithms already created to manage the gesture recognition process. The algorithmic logic that converts hand gestures into usable commands for the digital menu interface is contained in this endpoint.

4.3.5 Dynamic Interaction with Website

Flask's seamless integration allows users to dynamically interact with a website's user interface using simple hand gestures. Using the features of MediaPipe's gesture recognition systems, created by manually analyzing landmarks, predefined gestures such as page navigation and food selection are detected. Later, Flask interprets these gestures and triggers the corresponding actions in the website's user interface. This facilitates customers in familiarizing themselves with the culinary selections, seamlessly adding items to their shopping cart, and navigating the menu with exceptional ease and efficiency.

4.3.6 Algorithm Validation

To guarantee its resilience and dependability in a variety of situations, the proposed algorithm is rigorously validated. The algorithm's accuracy and efficiency are assessed by extensive testing protocols that include different hand positions, lighting situations, and user interactions. The purpose of this validation method is to verify the algorithm's effectiveness in practical situations and guarantee that the digital menu interface functions flawlessly.

To summarize, the Algorithm Development phase combines advanced computer vision techniques with web development paradigms to produce an immersive and intuitive dining experience. By seamlessly combining gesture recognition technology with a digital restaurant menu online, this research work transforms user interaction in the gastronomic arena. Flask's flawless integration with the HTML, CSS, and JavaScript website demonstrates the synergy between backend processing and frontend presentation, resulting in a transformational digital eating experience.

V. RESULTS AND DISCUSSION

The experimental phase of our research involved rigorous testing to evaluate the performance and accuracy of the hand gesture recognition algorithm in real-world scenarios, especially when ordering food. Our main goal was to carefully evaluate the algorithm's ability to accurately detect and interpret hand gestures, which were central to the smooth interaction of food scenarios.

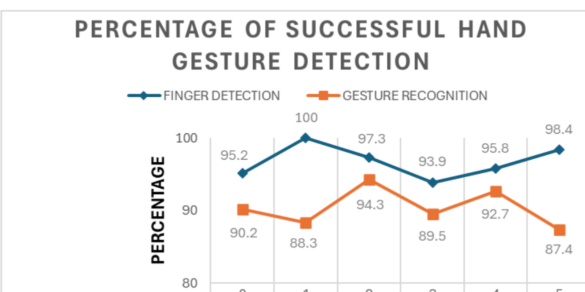
During extensive testing, we consistently observed strong and reliable results in both finger counting and gesture recognition exercises. The algorithm showed considerable ability in accurately distinguishing and cataloging fingers, achieving commendable success across the one- to five-finger spectrum. These results highlighted the flexibility and efficiency of the algorithm in capturing subtle variations in hand movements, which is crucial for accurate gesture interpretation. A Deep learning-based multi-modal approach also uses RGB and skeleton sequences for human activity recognition [21]. In addition, extensive evaluations were performed to measure the adaptability of the algorithm to various environmental factors and operational parameters. Evaluations dealing with changes in environmental noise levels, differences in light intensity, differences in hand size, and the presence of accessories had given encouraging results that highlighted the versatility and robustness of the algorithm in various conditions. This robust performance increases its practical applicability in real-world implementations, where environmental dynamics can vary unpredictably.

Along with independently evaluating the performance of the algorithm, we seamlessly integrate it into a web-based application adapted for digital food ordering. Using a combination of Flask, HTML, CSS, and JavaScript, the application provides an intuitive interface for users to navigate menu items and place orders using hand gestures. This integration not only demonstrates the algorithm's seamless compatibility with web-based technologies but also highlights its practical utility in improving the user experience. Offering a new and interactive way to select and order food, the app exemplifies the transformative potential of gesture recognition technology to redefine traditional dining paradigms.

Furthermore, user feedback and usability testing were conducted to optimize user interface design and user interaction. Iterative design improvements had been made based on user feedback and observations, resulting in a user-friendly and intuitive API. Incorporating user-centered design principles ensures that the application meets the needs and preferences of target users, improving overall usability and satisfaction. Introduction of many datasets like the Human Activity Recognition Trondheim dataset (HARTH), which is a publicly available dataset with professionally annotated labels and fixed Sensor placements, facilitating objective comparison between machine learning models for free-living human activity recognition [5] [22].

Overall, extensive testing and integration of the hand gesture recognition algorithm into a web-based application demonstrates the algorithm's robustness, adaptability, and practical utility in real-world scenarios. By providing users with a smooth and intuitive way to interact with digital user interfaces, our research contributed to the development of gesture recognition technology and its transformative potential to improve user experience in various fields.

The result of multiple attempts was averaged based on the number of successful hand gesture recognition for each finger count. It is observed from Figure 11 that the percentage of successful hand gesture recognition was performed with the help of our 5 group members.



	0	1	2	3	4	5
FINGER DETECTION	95.2	100	97.3	93.9	95.8	98.4
GESTURE RECOGNITION	90.2	88.3	94.3	89.5	92.7	87.4

Fig. 11. Showcasing line chart representing the percentage success of hand gesture detection

Through the implementation of a complex system for finger detection and hand gesture recognition as well as the use of powerful computer vision techniques and machine learning algorithms, our research has come up with a solution for this problem. This shows that the 96.766% of finger detection and 90.4% of gesture recognition accuracy rates reflect the effectiveness of our approach. The system determines fingertip position and hand landmarks configuration and then it accurately follows hand movements, thus providing the precision of finger detection. This is a very important matter because this approach is necessary for application that supports fine-hand gesture identification such as sign language translations and virtual touch interfaces.

The graphical representation of the percentage of successful hand gesture recognition was a visual proof of the effectiveness of our algorithm. With each data point carefully plotted, the graph vividly illustrates how the algorithm performs under different scenarios and conditions. From ambient noise levels to variations in light intensity, the diagram summarizes the adaptability and flexibility of the algorithm, providing valuable insight into its real-world applicability. As we delve into the nuances of motion detection, this diagram serves as a guiding sign, lighting the way for further refinement and optimization.

Recent research has highlighted several of the shortcomings of previous techniques for hand gesture identification. Xia et al. [23] demonstrated the computational expenses of using Convolutional Neural Network (CNN)--based approaches, which, despite their great accuracy, were sluggish to process and unsuitable for real-time applications. Another example was Xu et al. [16], who used a deep learning strategy to achieve high-precision findings but at a significant resource cost, which is unsuitable in resource-constrained applications. Furthermore, the user experience in gesture recognition systems is extremely important. The system's efficiency, which is dependent on easy setup and accessibility, is now hampered by the fact that many existing systems lack user-centric designs.

In contrast, our technique overcomes the constraints of either accurate or speedy prediction by combining them. We developed a web application that combines machine learning and web development techniques, leveraging high-end image processing to achieve high accuracy and real-time action detection with minimal hardware requirements. Furthermore, our solution prioritizes user experience; it interfaces with existing web platforms, eliminating the need for additional equipment, and can be simply inserted and played during setup procedures. Based on recent research findings and answers to recognized challenges, our idea represents a significant advancement in the gesture recognition technology field. To achieve a thorough approach, the technical and usability components of the system were considered. The existing paradigms in restaurant food ordering systems were reworked, among other things, to improve the user experience and promote wider usage of gesture recognition technology.

VI. CONCLUSION AND FUTURE WORK

Using a touchless interface that removes hygienic issues and physical interactions, our gesture-based food ordering system represents a revolutionary leap forward in dining experiences. The proposed solution transforms consumer interactions with ordering procedures, altering the constraints of conventional approaches and resolving touch-based interface concerns. It also establishes a new benchmark for ease and hygiene in restaurants. While developing intuitive interfaces for gesture-based food ordering systems, the assessment of digital menu types highlights the dynamic nature of menu presentation. A sophisticated hand gesture recognition system that demonstrates the resilience and versatility of the proposed work was put into practice as the proposed product. Redefining eating paradigms and improving user experiences are only two of the ways that the algorithm's ability to recognize gestures accurately across a variety of contexts is highlighted by its user-centered design approach.

This work has opened the door for more developments in gesture detection technology and offers insightful information about the larger field of human-computer interaction. The extensive research indicates a commitment to pushing the boundaries of technology by addressing issues, providing useful implementations, and setting the foundation for the next advancements.

REFERENCES

[1]. Jalal, A., Kamal, S., & Kim, D. (2016). Human depth sensors-based activity recognition using spatiotemporal features and hidden markov model for smart environments. *Journal of computer networks and communications*, *2016*.

[2]. Al Machot, F., & Mayr, H. C. (2016, June). Improving human activity recognition by smart windowing and spatio-temporal feature analysis. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (pp. 1-7).

[3]. Prati, A., Shan, C., & Wang, K. I. K. (2019). Sensors, vision and networks: From video surveillance to activity recognition and health monitoring. *Journal of Ambient Intelligence and Smart Environments*, *11*(1), 5-22.

[4]. Mukherjee, S., Anvitha, L., & Lahari, T. M. (2020). Human activity recognition in RGB-D videos by dynamic images. *Multimedia Tools and Applications*, *79*(27), 19787-19801.

[5]. Saleem, G., Bajwa, U. I., & Raza, R. H. (2023). Toward human activity recognition: a survey. *Neural Computing and Applications*, *35*(5), 4145-4182.

[6]. Hussain, Z., Sheng, M., & Zhang, W. E. (2019). Different approaches for human activity recognition: A survey. *arXiv preprint arXiv:1906.05074*.

[7]. Ferrari, A., Micucci, D., Mobilio, M., & Napoletano, P. (2020). On the personalization of classification models for human activity recognition. *IEEE Access*, *8*, 32066-32079.

[8]. [8] Mohamed, N., Mustafa, M. B., & Jomhari, N. (2021). A review of the hand gesture recognition system: Current progress and future directions. *ieee access*, *9*, 157422-157436.

[9]. Oudah, M., Al-Naji, A., & Chahl, J. (2020). Hand gesture recognition based on computer vision: a review of techniques. *journal of Imaging*, *6*(8), 73.

[10]. Guo, L., Lu, Z., & Yao, L. (2021). Human-machine interaction sensing technology based on hand gesture recognition: A review. *IEEE Transactions on Human-Machine Systems*, *51*(4), 300-309.

[11]. Al-Hammadi, M., Muhammad, G., Abdul, W., Alsulaiman, M., Bencherif, M. A., & Mekhtiche, M. A. (2020). Hand gesture recognition for sign language using 3DCNN. *IEEE access*, *8*, 79491-79509.

[12]. Parvathy, P., Subramaniam, K., Prasanna Venkatesan, G. K. D., Karthikaikumar, P., Varghese, J., & Jayasankar, T. (2021). Development of hand gesture recognition system using machine learning. *Journal of Ambient Intelligence and Humanized Computing*, *12*, 6793-6800.

[13]. Mujahid, A., Awan, M. J., Yasin, A., Mohammed, M. A., Damaševičius, R., Maskeliūnas, R., & Abdulkareem, K. H. (2021). Real-time hand gesture recognition based on deep learning YOLOv3 model. *Applied Sciences*, *11*(9), 4164.

[14]. Zhang, W., Wang, J., & Lan, F. (2020). Dynamic hand gesture recognition based on short-term sampling neural networks. *IEEE/CAA Journal of Automatica Sinica*, *8*(1), 110-120.

[15]. Sharma, A., Mittal, A., Singh, S., & Awatramani, V. (2020). Hand gesture recognition using image processing and feature extraction techniques. *Procedia Computer Science*, *173*, 181-190.

[16]. Xu, C., Chai, D., He, J., Zhang, X., & Duan, S. (2019). InnoHAR: A deep neural network for complex human activity recognition. *Ieee Access*, *7*, 9893-9902.

[17]. Bevilacqua, A., MacDonald, K., Rangarej, A., Widjaya, V., Caulfield, B., & Kechadi, T. (2019). Human activity recognition with convolutional neural networks. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2018, Dublin, Ireland, September 10–14, 2018, Proceedings, Part III 18* (pp. 541-552). Springer International Publishing.

[18]. Wan, S., Qi, L., Xu, X., Tong, C., & Gu, Z. (2020). Deep learning models for real-time human activity recognition with smartphones. *Mobile Networks and Applications*, *25*(2), 743-755.

[19]. Ramanujam, E., Perumal, T., & Padmavathi, S. (2021). Human activity recognition with smartphone and wearable sensors using deep learning techniques: A review. *IEEE Sensors Journal*, *21*(12), 13029-13040.

[20]. Kaur, S., Shukla, H. K., Pal, R. K., Yadav, N., & Singh, S. (2022). Human Activity Recognition.

[21]. Verma, P., Sah, A., & Srivastava, R. (2020). Deep learning-based multi-modal approach using RGB and skeleton sequences for human activity recognition. *Multimedia Systems*, *26*(6), 671-685.

[22]. Logacjov, A., Bach, K., Kongsvold, A., Bårdstu, H. B., & Mork, P. J. (2021). HARTH: a human activity recognition dataset for machine learning. *Sensors*, *21*(23), 7853.

[23]. Xia, K., Huang, J., & Wang, H. (2020). LSTM-CNN architecture for human activity recognition. *IEEE Access*, *8*, 56855-56866.

Ashlin Joel was born in Hyderabad, India, in 2004. He has done his schooling from St. Patrick's High School, Secunderabad, Telangana, India. He is currently pursuing Computer Science Engineering from Vellore Institute of Technology, Vellore, Tamil Nadu, India. His areas of interest include Full Stack Development, Interface Design and Deployment, and Cloud Computing.

Navamani T M received the M.E. (C.S.E.) and Ph.D. degrees from Anna University, Chennai. She is currently a Senior Associate Professor at the School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Vellore, India. She has 26 years of teaching experience. She has published more than 40 papers in international and national journals, conferences, and book chapters. Her research interests include data sciences, human–computer interaction, wireless ad hoc networks, network security, and mobile computing. She is a Life Member of the Indian Society for Technical Education.

Archita Sharma was born in Delhi, India, in 2002. She has done her schooling from Modern Delhi Public School, Faridabad, India. She is currently pursuing Computer Science Engineering from Vellore Institute of Technology, Tamil Nadu, India. Her research interests include App development, Artificial Intelligence & Web development.

Vansh Harkut was born in Hyderabad, India, in 2002. He received his primary education from the Future Kids School, Nanakramguda, Hyderabad, India, and his Secondary education from FIITJEE Junior College, Narayanguda, Hyderabad, India. He is currently pursuing the B.Tech. degree in computer science and engineering with the Vellore Institute of Technology, Tamil Nadu, India. His research interests include Artificial Intelligence, Machine Learning, and Computer Vision.

Mansi Saxena was born in Indore, India in 2003. She has done her schooling at Carmel Convent Sr. Sec. School B.H.E.L Bhopal. And is currently pursuing Computer science engineering from Vellore Institute of Technology, Vellore, India. Her areas of interest include Cloud computing, web development, and data security.

Arnav Goenka was born in Delhi, India in 2003. He has done his schooling from G.D Goenka Public School, Delhi, and is currently pursuing Computer Science Engineering with Specialization in Bioinformatics from Vellore Institute of Technology, Vellore, Tamil Nadu, India. His research interests include Artificial Intelligence, Machine Learning, and Data Analytics.