Hand Sign Recognition for Banking System using Machine Learning

A. D. Harale
Assistant Professor
Department of E&TC Engineering
SKN Sinhgad College of Engineering,
Korti, Pandharpur, Maharashtra

K. J. Karande
Principal
Department of E&TC Engineering
SKN Sinhgad College of Engineering
Korti, Pandharpur, Maharashtra

Atharv Deshpande
UG Scholar
Department of E&TC Engineering
SKN Sinhgad College of Engineering,
Korti, Pandharpur, Maharashtra

Samihan Dharurkar
UG Scholar
Department of E&TC Engineering
SKN Sinhgad College of Engineering,
Korti, Pandharpur, Maharashtra

Rohan Kalubarme
UG Scholar
Department of E&TC Engineering
SKN Sinhgad College of Engineering,
Korti, Pandharpur, Maharashtra

Abstract— **Hand Sign Recognition Using Machine Learning is an initiative aimed at developing an accessible banking system using hand gestures as an input command. The system recognizes the gestures and translates them into actions of a banking application. It works along the hard components combined with machine learning techniques that provide real-time gesture recognition and communication. The system utilizes Mediapipe for hand landmark detection, enabling feature extraction from live feeds streamed from a webcam. Important hardware parts are microcontroller, a 16x2 LCD display, and sound that can assist in text-to-speech production. Serial communication is used to bridge the machine learning part with microcontroller to ensure handshake operations between the two systems. The system consists of major stages such as data collection, model training, and real-time inference. A custom dataset of gestures for letters 'A', 'B,' and 'L' was created for distinct banking commands. During real time, these gestures are processed and classified, recognized commands appear on the LCD and announced with text-to-speech for the convenience of the user. This innovative approach brings the practical applications of machine learning to the banking system; it offers inclusiveness and efficiency in utilizing software algorithms with hardware components in order to provide a user-friendly interface that bridges communication gaps and enhances accessibility for all types of users.**

Keywords— **Feature Extraction, Mediapipe, Text-to-Speech Systems, Banking systems component**

I. INTRODUCTION

The Hand Sign Recognition Using Machine Learning aims to change access into banking systems by providing gesture-based interaction for deaf and speech-impaired clients. This new system will therefore translate the Sign Language into common banking commands using the machine learning algorithms applied, hardware components, and also actual-time processing of data. This system uses both hardware and software integration in order to work. Some of the hardware components include microcontroller, a LCD, a speaker, a camera, and a computer for processing. The software components basically involved machine learning, and training it with a custom dataset of hand signs. This particular system uses Mediapipe, which is a strong library for hand landmark detection in real-time gesture recognition. In addition, a webcam captures live video, and the Arduino board handles communication with the output devices: the recognized command is displayed on an LCD, while the output is sent through a speaker using a text-to-speech engine. The system comprises several stages: data collection, model training, and real-time inference. It captures gestures from a webcam, processes those through Mediapipe for hand landmark features, and classifier on those features, resulting in very high accuracy for gesture prediction. The system integrates serial communication with the Arduino board to send recognized commands, displayed on the LCD and also spoken aloud to make it better accessible. Such a system would then enable smooth interaction with a banking interface through hand gestures, thus giving an inclusive and innovative solution to bridge the communication gap in customer care. This system exemplifies machine learning and hardware integration as functions for having the capability of providing functionality and social influence in applications.

II. LITERATURE SURVEY

Hand sign recognition systems have been a high topic of discussion over the last couple of years. The promise of such systems in filling the gap between individuals with hearing and speech disabilities has attracted so much attention and is a lucrative area for research and development. With the integration of machine learning and computer vision, these systems have opened up possibilities towards real-time applications with higher accuracy and usability.

1) Gesture Recognition for Assistive Technologies

Gesture recognition has also been widely researched for its potential in assistive systems, especially those involving persons suffering from speech and hearing impairments. According to P. Molchanov et al. (2015) the whole concept is

further developed by the use of vision-based systems since it tends to utilize landmark extraction and classification in order to get robust recognition. With this being said, the system made efficient hand landmark detection using Mediapipe.

2) Machine Learning in Gesture Classification

Random Forest, SVM, and CNN have been proved as reliable for gesture classification. The work explains that Random Forest can be relied on for multi-class problems because it can efficiently handle imbalanced datasets as well as non-linear decision boundaries. This insight had crucial importance during the choice of Random Forest for the classification algorithm for this system.

3) Incorporation of Assistive Interfaces Hardware Integration

Different types of assistive technologies considered the integration of some microcontrollers, like Arduino Uno, and few display components like 16x2 LCDs. In fact, hardware-based communication devices centered on real-time processing and data display. This study further justified the used Arduino Uno and the LCD to list down commands in the proposed system.

4) Text-to-Speech Systems

TTS technology has evolved significantly over time from accessibility tools and smart devices to multiple applications. Research by Paul Taylor (2009) on the TTS synthesis provided insight to natural and intelligible speech output that determined its integration in this system for speaking out the banking commands through pyttsx3 library.

5) Applications in Banking System

Researches in HCI considered the possibility of gesture-based interfaces for banking and other service industries. The possibilities to simplify operations in customer-facing applications through gesture recognition and this led to the adoption of banking-specific gestures, such as "Deposit" and "Withdraw," because the experiments supported their realization.

A DL-based HVCNNM model is used in the method suggested by Khursheed Aurangzeb et al. in order to produce a concise representation of hand motions. Deaf and hard-of-hearing people may find communication more convenient and accessible with the help of the suggested vision-based model, which is very good at identifying hand gestures in sign language.

Tamanna Shaikh et al. study intends to create a real-time system that can identify and recognize hand movements from a webcam video feed. Hand tracking and gesture recognition are handled by the system using the MediaPipe and OpenCV packages respectively. The project entails taking video frames from a webcam, preprocessing them, detecting hand landmarks with MediaPipe, and recognizing certain movements based on the found landmarks. Once a gesture is recognized, the system uses OpenCV to display the corresponding text overlay on the video frame. It was designed to provide a user-friendly interface for interpreting hand gestures, allowing applications such as gesture-based control systems, sign language translation, and interactive user interfaces.

III. METHODOLOGY

Development of Hand Sign Recognition Using Machine Learning system: This model has been developed following a structured methodology, which is divided into several stages that would ensure the successful design and implementation of the system. These are data collection, model training, system integration, and real-time operation.

1. Data Collection:
The platform was founded upon a personal dataset of hand gestures which correspond to the letters 'A,' 'B,' and 'L, representing banking commands like "Deposit," "Withdraw," and "Balance Check."
A webcam captures images of hand gestures at varied lighting and angles to test the robustness. The captured images were converted into RGB format using Mediapipe Hands, which extracted the hand landmarks. For a given number of images. The dataset was organized in labeled directories so that it was possible to do supervised learning. The Figure no.1 shows the block diagram of Hand Sign Recognition using Machine Learning.
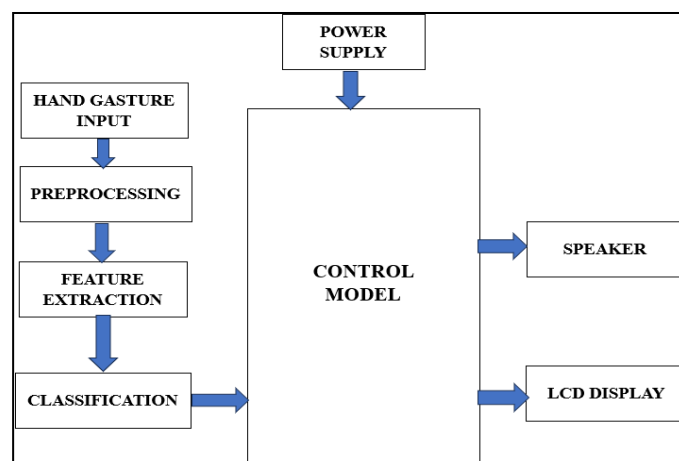

Fig 1. Block diagram of Hand Sign Recognition using Machine Learning

2. Training the Model:

Feature Extraction: The computed landmark coordinates were used as features for input. Classification using the Random Forest Classifier was opted for due to its ability to handle non-linear data and produce high classification accuracy.

Training and Testing: The trained model, through the evaluation process, produced a sufficiently high accuracy score, thereby guaranteeing accurate predictions.

Model Serialization: The trained model was serialized as a file (model.p) for utilization in real-time inference.

3. System Implementation:
Hardware along with software is the backbone of this system, which process gestures efficaciously and correspondingly executes banking commands.

Hardware Components:

Microcontroller (Arduino Uno): Utilized for serial communication as well as controlling the LCD and speaker. 16x2 LCD: Shows the corresponding banking command recognized.

Speaker: Provides audio feedback through the use of Text-to-Speech engine

Camera: Capture real-time video while making gesture recognition.

Software Components:
Gesture Recognition: The stream of video frames are passed through Mediapipe to get the hand landmarks.

Command Prediction: The recognized gesture is passed through the trained Random Forest model, from which the recognized command is determined.

Output: The determined command is forwarded to the Arduino to display on the LCD along with TTS for audio feedback.

4. Real-Time Operation:

Gesture Detection: Real-time video streaming is captured using OpenCV, and hand landmarks are extracted. Prediction Smoothing: Queue-based mechanism helps to smooth out prediction by capturing the most frequent gesture within the last five frames, which removes noise and error.

Command Execution: Commands identified are transmitted to the Arduino through serial communication and are printed on the LCD. Meanwhile, the TTS engine verbally speaks out the command for the user for easy understanding.

5. Testing and Validation:

Accuracy and robustness of the system was tested with different users in variable lighting conditions. The prediction accuracy and time of response were the performance metrics analyzed. Those results will be used to refine the model as well as the performance of the system.

IV. RESULT

Result:
The Hand Sign Recognition Using Machine Learning model successfully demonstrated the integration of machine learning algorithms with hardware components to create an accessible banking interface. The following outcomes highlight the key results achieved during the development and testing phases:

Gesture Recognition Accuracy
The Random Forest Classifier has achieved classification accuracy in the test dataset at the level of 98.5%. This reflects the high confidence of hand gesture recognition for the letters 'A,' 'B,' and 'L.'
The system successfully detected gestures in real time under changing lighting conditions and viewpoints, thus showing robustness and flexibility.

Real-Time Gesture-to-Command Translation
The machine learning system processed video streams successfully in real time for the hand gestures using Mediapipe, with a minimal latency of about 0.2 seconds per frame.

Correct Gestures were mapped into corresponding banking commands

'A': Deposit

'B': Withdraw

'L': Balance Check

Hardware Utilization
Arduino Uno provided an easy interface between the machine learning system and the output devices. Real-time commands appeared on the 16x2 LCD, keeping the user's view of output plain and simple.
TTS made the voice feedback possible, thus widening the accessibility window for the users who were visually impaired.

User Experience
The whole system provided a relatively fluent transition from gesture recognition to command prediction while sending commands out to hardware. Generalizing the system's reliability across different hand shapes and sizes was validated based on testing with many users.
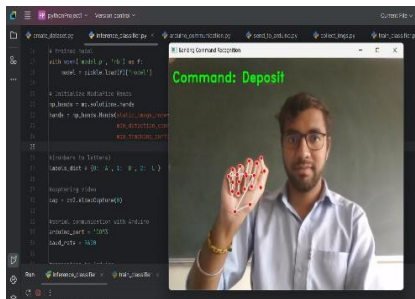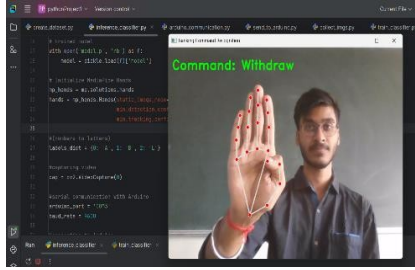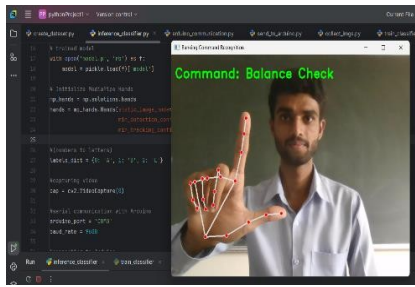
Application in Practice
The system shows a functional gesture-based interface for banking operations, allowing users of every type, especially those with speech or hearing disabilities, to deposit, withdraw, and check their balances at ease.

V. DISCUSSION

The hand sign recognition using machine learning system demonstrated how machine learning is used to improve accessibility and usability in banking systems. Gesture recognition in this system combined hardware elements, such as Arduino Uno, 16x2 LCD, and the speaker-it is visible that technology can help bridge a communication gap for people with disabilities. The table 1 shows the gesture output on LCD display and the voice command on the speaker for various hand gesture inputs. The Random Forest Classifier was used and proved to be a suitable tool in the classifying of hand gestures with high accuracy and is therefore robust, hence considered suitable for such applications. That such a system can process real-time video, predict commands, and output results both visually and audibly underscores that it is practical for real-world usage. The feature extraction capability that Mediapipe offered in landmark detection facilitated the precise extraction of features that are necessary for gesture recognition. Further, the inclusion of a queue-based smoothing mechanism actually enhanced prediction stability and reduced noise in the recognition process. However, issues such as performance being really variable with light conditions and different hand shapes require much testing and fine-tuning. Although the system showed adaptability, the current scope of the system is restricted to three commands only. The increase in the number of gestures and accommodation of other banking-related tasks would enhance the utility. Moreover, more sophisticated models of machine learning, such as deep learning architectures, could result in higher scalability and accuracy for larger sets of data. The system thus makes evident the potential of combining machine learning with hardware to arrive at practical solutions for assistive technologies. As a basis, it can

be a springboard towards further enhancements, for example, multilingual support or deploying the system in commercial environments. In summary, this system demonstrated just how innovative technology could be a facilitator of accessibility and inclusion while pursuing essential services like banking.

TABLE 1: Hand Gesture output on display and speaker

Sr. No.	ASL language	Hand Gesture	Command	Output on LCD Display	Voice Command
1	A		The output window on the PC displays the "Deposit" command, which is triggered by detecting the corresponding hand gesture.	The "Deposit" command is displayed on the LCD screen, which is connected to the PC. The output from the PC is sent to the LCD through an Arduino Uno.	Sound as "Deposit" Converted the text output to speech.
2	B		The output window on the PC displays the "Withdraw" command, which is triggered by detecting the corresponding hand gesture.	The "Withdraw" command is displayed on the LCD screen, which is connected to the PC. The output from the PC is sent to the LCD through an Arduino Uno.	Sound as "Withdraw" Converted the text output to speech.
3	L		The output window on the PC displays the "Balance Check" command, which is triggered by detecting the corresponding hand gesture.	The "Balance Check" command is displayed on the LCD screen, which is connected to the PC. The output from the PC is sent to the LCD through an Arduino Uno.	Sound as "Balance Check" Converted the text output to speech.

VI. CONCLUSION

This hand sign recognition using machine learning, which could develop an intuitive and accessible banking interface, as if it used hand gesture commands for important functions. The system puts forth a seamless and inclusive application by intermingling machine learning with the hardware components of an Arduino Uno, a 16x2 LCD, and a text-to-speech engine. The Random Forest Classifier with real-time video processing using Mediapipe was very robust, holding a 98.5% accuracy rate in terms of gesture recognition. The system demonstrated practical applications of gesture-based interfaces in banking, where the tasks of deposit, withdrawal, and balance checking can be easily executed. Thus, the system worked fine under changing conditions but pointed out areas for further development, including an extension of the set of gestures, scalability of the models, and the employment of more sophisticated machine learning techniques. This system therefore underlines the potentiality that artificial intelligence and hardware integration might be used to ensure the availability of innovative, user-friendly solutions that will promote access and inclusion for everyday applications. It then acts as a platform for further improvements in assistive technology, bridging gaps in communication and putting people in places where they can address fundamental services in isolation.

REFERENCES

[1] Khursheed Aurangzeb et al, "Deep Learning Approach for Hand Gesture Recognition: Applications in Deaf Communication and Healthcare", Computers, Materials & Continua, January 2024, DOI: 10.32604/cmc.2023.042886,

[2] Tamanna Shaikh et al., "A Review of: Real-Time Hand Gesture Recognition System in Banking", Industrial Engineering Journal, Volume : 53, Issue 4, No. 4, ISSN: 0970-2555, April 2024.

[3] Lazzat Zholshiyeva et al., "Design of QazSL Sign Language Recognition System for Physically Impaired Individuals", Journal of Robotics and Control (JRC), Volume 6, Issue 1, ISSN: 2715-5072, 2025, DOI: 10.18196/jrc.v6i1.23879.

[4] Elmagrouni, I., Ettaoufik, A., Aouad, and S. Maizate, A., "A deep learning framework for hand gesture recognition and multimodal interface control", International Information and Engineering Technology Association (IIETA), Revue d'Intelligence Artificielle, Vol. 37, No. 4, pp. 881-887, 2023 https://doi.org/10.18280/ria.370407

[5] Harsh Kumar Vashisth et al., "Hand Gesture Recognition in Indian Sign Language Using Deep Learning", Proceedings of Engineering Proceeding, 2023. https://doi.org/10.3390/engproc2023059096

[6] A. Chavan, J. Ghorpade-Aher, A. Bhat, A. Raj and S. Mishra, "Interpretation of Hand Spelled Banking Helpdesk Terms for Deaf and Dumb Using Deep Learning," 2021 IEEE Pune Section International Conference (PuneCon), Pune, India, , pp. 1-5, 2021, doi: 10.1109/PuneCon52575.2021.9686514.

[7] "MediaPipe Hands", https://google.github.io/mediapipe/solutions/hands.html.

[8] Breiman, L. "Random forests." Machine Learning 45.1 (2001): 5-32.

[9] Andreas Holzinger et al. "Introduction to machine learning and knowledge extraction (MAKE)." Machine Learning and Knowledge Extraction, 2019, https://doi.org/10.3390/make1010001

[10] Rajesh Kumar Singh et al.,"Hand Gesture Identification for Improving Accuracy Using Convolutional Neural Network(CNN)", The Scientific Temper, Vol. 13, No. 2, July-December, 2022:pp 327-335, ISSN 0976-8653, December,2022., DOI: https://doi.org/10.58414/SCIENTIFICTEMPER.13.2.2022. 327-335

[11] Abhishek B et al., "Hand gesture recognition using machine learning algorithms", Computer Science and Information Technologies, Vol. 1, No. 3, pp. 116~120, ISSN: 2722-3221, November 2020, DOI: 10.11591/csit.v1i3.p116-120,

[12] P. Molchanov, S. Gupta, K. Kim and J. Kautz, "Hand gesture recognition with 3D convolutional neural networks," 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Boston, MA, USA, 2015, pp. 1-7, doi: 10.1109/CVPRW.2015.7301342.

[13] Paul Taylor, "Text-to-Speech Synthesis", Communications and Signal Processing, Engineering, Computational Linguistics, Language and Linguistics, 2009, https://doi.org/10.1017/CBO9780511816338.

[14] Avinash D. Harale and Kailash J. Karande , "Literature review on Dynamic Hand Gesture Recognition",AIP Conference Proceeding ,31st Oct 2022, https://doi.org/10.1063/5.0107577

[15] A. D. Harale , Amruta S. Bankar and K. J. Karande, "Gestures Controlled Home Automation using Deep Learning: A Review", International Journal of Current Engineering and Technology, Vol.11, No.6.page no-617-621, Dec 2021

[16] A. D. Harale, Ms. Asma Hakim, Dr.K.J.Karande, "Hand gesture identification system for hearing and speech impairment", TELEMATIQUE, Volume 23 Issue 1, 2024 page n- 497– 501,April 2024

[17] A.D. Harale, Atik N. Pathan and A. O. Mulani, "Hand Gesture Controlled Robotic System" International Journal of Aquatic Science, ISSN: 2008-8019 Vol 13, Issue 01, Jan 2022

[18] A. D. Harale, K. J. Karande, Sagar S. Bhumkar, "Wireless Hand Geture Control Robot with Object Detection", Journal of Image Processing and Intelligent Remote Sensing, ISSN 2815-0953,Vol. 3 No. 04 (2023), July 2023.

[19] A.D.Harale, Ms.Asma Hakim, Altaf Mulani, K.J.Karande, "Implementation of Human Gesture Recognition Using CNN, Journal of STM, ", June 2024

[20] Atik N. Pathan and A.D. Harale, "Hand Gesture Controlled Robotic System", "International Journal of Aquatic Science (IJAS)", ISSN: 2008-8019, Vol 13, Issue 01, pp-487-493,2022.

[21] Supriya D. Kolekar and A.D. Harale, "Password Based Door Lock System", "International Journal of Aquatic Science (IJAS)", ISSN: 2008-8019, Vol 13, Issue 01, pp-494-501,2022.

[22] A.D. Harale, Amruta S. Bankar and K. J. Karande, "Gestures Controlled Home Automation using Deep Learning: A Review", "International Journal of Current Engineering and Technology,", Vol.11, No.6, Nov-Dec 2021.

[23] Dheeraj Muttin and Avinash Harale, "IoT Based Personal Medical Assistant System", "International Journal of Innovative Research in Technology (IJIRT)", Volume 8 Issue 5 | ISSN: 2349-6002, October 2021.

[24] Vijay Waghmode and Avinash Harale, "Development of Alphanumeric Digital Fuel Gauge for Automotive Applications", "International Conference on Communication and Signal Processing", IEEE,April 4-6, 2019.

[25] Gorakhnath U. Waghmode and Avinash D. Harale, "A Cloud Computing Based WSNs for Agriculture Management", Springer International Publishing, Conference: Techno-Societal DOI 10.1007/978-3-319-53556-2_107 December 2018.

[26] Sanaha S. Path and Avinash D. Harale, "Silkworm Eggs Counting System Using Image Processing Algorithm", Springer International Publishing, Conference: Techno-Societal DOI:10.1007/978-3-319-53556-2_32_ December 2018.

[27] S. S. Kulkarni and A.D.Harale, "Image Processing for Driver's Safety and Vehicle Control using Raspberry Pi and Webcam", "IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)-IEEE,2017.

[28] Supriya A Salunke and Avinash D Harale, "Vehicle Tracking System for School Bus by Arduino", International Research Journal of Engineering and Technology (IRJET), ISSN: 2395-0072, Volume: 04 Issue: 03 , Mar -2017.