International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Gesture PPT: AI based Two-Stream CNN-based Hand Gesture Recognition System for PowerPoint Control

*** K.Dhavana [1], Mr.J.Jayapandian [2]**

[1]Master of Computer Applications, Krishnasamy College of Engineering &Technology, Cuddalore,India
[2]MCA.,M.Phil.,(Ph.D).,Associate Professor, Master of Computer Applications, Krishnasamy College of Engineering &Technology, Cuddalore,India.

ABSTRACT :

In today's digital world Presentation using a slideshow is an effective and attractive way that helps speakers to convey information and convince the audience. There are ways to control slides with devices like mouse, keyboard, or laser pointer, etc Traditional input devices for controlling presentations, such as keyboards or remote controls, may not always provide the desired level of user convenience, particularly in scenarios where physical proximity or direct manipulation is crucial. Presenters often face challenges when navigating slides or interacting with content during a presentation. The incorporation of hand gesture recognition technology aims to address these challenges by providing a more fluid and interactive means of controlling presentations. This project introduces a Hand Gesture Recognition-based PowerPoint Control System utilizing a Two-Stream Convolutional Neural Network (CNN). The system aims to provide an intuitive and efficient interface for controlling Microsoft PowerPoint presentations through real-time interpretation of hand gestures. The Two-Stream CNN architecture is employed to concurrently process spatial and temporal features, capturing both the static and dynamic aspects of hand movements. The spatial stream of the network focuses on analyzing still frames, identifying key points and spatial relationships within the hand gestures. Simultaneously, the temporal stream processes optical flow information between consecutive video frames, capturing the dynamic motion patterns associated with different gestures. Integrating these streams enables the system to comprehensively understand and interpret a diverse range of hand gestures. To train and evaluate the system, a curated dataset featuring various hand gestures for PowerPoint control functions, such as slide navigation, start, and stop, is employed. Experimental results demonstrate the system's efficacy in accurately recognizing a diverse set of hand gestures, providing a responsive and user-friendly interface for PowerPoint control.

Keywords: Hand Gesture Recognition-based PowerPoint Control ,Two-Stream CNN, Convolutional Neural Network (CNN).

INTRODUCTION :

Gestures, a form of non-verbal communication, encompass visible bodily actions like hand movements, facial expressions, or other body parts to convey specific messages, either alongside or instead of speech. Unlike non-verbal cues that don't communicate specific messages, gestures can convey a wide range of emotions and thoughts, such as approval, affection, contempt, or hostility, and are often used together with spoken language to enhance meaning. Some gestures, particularly hand gestures, can act like words themselves, having fixed meanings within a culture but varying significantly across different cultures and even among sub-communities within the same culture. This cultural specificity makes categorizing gestures challenging, especially as they evolve over time, such as the shift from the "call me" gesture mimicking a traditional phone to a flat palm signifying a smartphone. Traditional systems for controlling PowerPoint presentations, like keyboards, mice, and presentation clickers, present notable limitations. These devices require the presenter to stay close to the presenting device, restricting their movement and often interrupting their engagement with the audience due to the need to operate the control device manually.

Additionally, these methods can be imprecise and may suffer from lag or missed commands. Existing gesture recognition algorithms also struggle with accuracy, real-time performance, and adaptability to various environments. These issues underscore the necessity for a more intuitive and hands-free solution that leverages advanced technologies such as AI and deep learning to improve presentation control. Artificial intelligence (AI) encompasses systems that mimic human cognitive functions like problem-solving and learning, with applications in areas such as facial and speech recognition, decision-making, and translation. Machine learning, a subset of AI, involves algorithms that improve through experience. Deep learning, a specialized area within machine learning, uses neural networks with multiple layers to process data in a manner inspired by the human brain, allowing for the recognition of complex patterns in images, text, and sounds.

Convolutional neural networks (CNNs), a type of deep learning model, are particularly effective in analyzing visual data, distinguishing themselves through their ability to learn features from data via convolutional and pooling layers. In the context of video analysis, the two-stream network architecture is used to capture spatial and temporal components. This involves processing still frames and optical flow information through separate convolutional neural network (ConvNet) streams, which are then combined through a late fusion technique. The spatial stream extracts information about scenes and objects, while the temporal stream captures motion information through optical flow displacement. This architecture enhances the ability to recognize complex activities in videos by combining the strengths of both streams and passing the fused information through fully connected layers for classification. The aim of this project is to develop a gesture-based PowerPoint controller system that enhances the presentation experience by allowing users to control slides with hand gestures. This involves creating a web-based application with functionalities for slide navigation and presentation control via a user-friendly interface. A key component is the development of a robust hand gesture recognition system using computer vision techniques and deep learning models. By training a convolutional neural network on a dataset of hand gestures, the project seeks to achieve real-time and accurate gesture recognition, seamlessly integrating this capability into the PowerPoint controller application.

The project's scope includes developing a comprehensive solution for gesture-controlled PowerPoint presentations. This entails creating a web application that supports various presentation controls accessible through a web browser and implementing a hand gesture recognition system that accurately interprets gestures in real-time. The project also involves training and deploying a deep learning model for gesture recognition and integrating it with the PowerPoint controller to enable intuitive interactions. Additional functionalities include presentation management, user authentication, and access control to ensure data security and privacy, ultimately providing a seamless and interactive presentation experience.

LITERATURE SURVEY :

2.1. Deep Learning-Based Approach for Sign Language Gesture Recognition with Efficient Hand Gesture Representation

Author: Muneer Al-Hammadi; Ghulam Muhammad
Year:2020
Doi: 10.1109/ACCESS.2020.3032140
Problem
The paper addresses the challenge of dynamic hand gesture recognition, particularly focusing on sign language translation. It highlights the complexities involved in recognizing hand gestures, such as hand segmentation, local and global feature representation, and sequence modeling. The importance of accurate hand gesture recognition is emphasized, especially in touchless applications and for the hearing-impaired population. Objective Develop a novel system for dynamic hand gesture recognition that overcomes the challenges mentioned. Propose a system capable of accurately recognizing hand gestures in uncontrolled environments. Investigate the effectiveness of multiple deep learning architectures and techniques for addressing various aspects of hand gesture recognition. Evaluate the proposed system on a challenging dataset and compare its performance with state-of-the-art approaches. Methodology and Segmentation: Utilize the OpenPose framework for hand region detection and estimation. Feature Representation: Employ multiple deep learning architectures, including 3DCNN, MLP, and autoencoders, for capturing local hand shape features and global body configuration features. Sequence Feature Globalization and Recognition: Aggregate and globalize the extracted features using MLP and autoencoders, and employ the SoftMax function for classification. Training Optimization: Investigate domain adaptation techniques to reduce the training cost of the 3DCNN module. Evaluation: Assess the performance of the proposed system on a challenging dataset consisting of dynamic hand gestures performed in real-world setting Finding The proposed system outperforms state-of-the-art approaches in terms of recognition rate. It demonstrates effectiveness and robustness in recognizing dynamic hand gestures, even in uncontrolled environments. The integration of multiple deep learning architectures and techniques enables accurate representation and classification of hand gestures, particularly beneficial for sign language translation. The investigation into domain adaptation techniques for reducing training cost provides insights into optimizing deep learning models for hand gesture recognition.

2.2. A Novel Hybrid Deep Learning Architecture for Dynamic Hand Gesture Recognition

Author: David Richard Tom Hax; Pascal Penava
Year:2024
Doi: 10.1109/ACCESS.2024.3365274
Problem
The paper identifies several challenges and gaps in vision-based hand gesture recognition systems: Limited progress in continuous gesture recognition. Restriction of existing databases to controlled environments, hindering real-world applicability. Importance of addressing issues in data acquisition, data environment, and hand gesture representation for improved system performance. Objective The objective of the paper is to review and analyze the progress, challenges, and future directions of vision-based hand gesture recognition systems from 2014 to 2020.Specifically, the paper aims to: Identify key areas of focus in hand gesture recognition research. Assess recognition accuracy in signer-dependent and signer-independent scenarios. Evaluate the impact of data acquisition, data environment, and hand gesture representation on system performance. Highlight the need for less restrictive sign language databases and improved recognition in uncontrolled environments. Methodology The paper employs a systematic review methodology, extracting relevant articles from well-known online databases using selected keywords related to vision-based hand gesture recognition. Articles are selected based on their relevance to the research objectives and inclusion criteria. Key data such as recognition accuracy and critical aspects of hand gesture recognition (data acquisition, data environment, hand gesture representation) are extracted and analyzed. The review covers a seven-year period (2014-2020) to provide a comprehensive overview of research progress during this time frame. Finding The review highlights the active nature of vision-based hand gesture recognition research, with numerous studies conducted and published annually. Recognition accuracy for signer-dependent scenarios ranges from 69% to 98%, with an average accuracy of 88.8%, while signer-independent scenarios range from 48% to 97%, with an average accuracy of 78.2%. Critical aspects such as data acquisition, data environment, and hand gesture representation are identified as key areas of focus in the reviewed literature. The predominance of databases from restricted environments underscores the need for more diverse and less restrictive datasets to improve real-world applicability. The paper concludes that addressing issues in uncontrolled environment settings is essential to enhancing the system's ability to recognize hand gestures across different contexts and preparing it for practical real-life applications.

2.3. Hand Gesture Recognition with Flexible Capacitive Wristband Using Triplet Network in Inter-Day Applications

Author: Tiantong Wang; Yunbiao Zhao;
Year:2022
Doi: 10.1109/TNSRE.2022.3212705
Problem

Inter-day hand gesture recognition across multiple sessions and days presents challenges due to variations in hand positioning, pressure distribution, and environmental factors. Existing methods often struggle to maintain acceptable recognition accuracy over extended periods of use. This study aims to address these challenges by developing a novel approach that combines a flexible wristband with highly sensitive capacitive pressure sensing and advanced feature embedding techniques. Objective The objective of this study is to enhance inter-day hand gesture recognition accuracy using a wearable, flexible multi-channel capacitive wristband integrated with a triplet network for deep feature embedding. Methodology Wristband Design: A flexible wristband integrated with a highly sensitive capacitive pressure sensing array was developed. Experimental Setup: Seven hand gestures were selected, and inter-day experiments were conducted over five consecutive days, with three sessions each day. Five healthy subjects participated, with the wristband being doffed and re-donned between sessions. Data Collection: Pressure distributions around the wrist were captured using the capacitive array, treating the outputs as frames of images. Triplet Network Implementation: A triplet network architecture was utilized for deep feature embedding. This network consisted of three identical CNN structures trained with triplet loss. Performance Evaluation: Recognition accuracy was measured for each session and subject, comparing the performance of the triplet network with a conventional CNN trained with softmax-cross-entropy loss. The influence of capacitive array size on recognition results was also investigated. Finding Triplet Network Performance: The triplet network achieved an average recognition accuracy of 91.98%, outperforming the conventional CNN by 7.33% ($p < 0.05$). Influence of Capacitive Array Size: The full-size capacitive array (4×8) yielded significantly better results compared to down-sampled arrays ($p < 0.05$), indicating the importance of array size in hand gesture recognition accuracy. Feasibility of Approach: The study demonstrates the feasibility of improving inter-day hand gesture recognition performance using a wearable, flexible multi-channel capacitive wristband and a triplet network for deep feature embedding. This approach offers promising insights into addressing the challenges of prolonged usage of human-machine interfaces for hand gesture recognition.

2.4. TraHGR:Transformer for Hand Gesture Recognition via Electromyography

AUTHOR: Soheil Zabihi; Elahe Rahimian;
Year:2023
Doi: 10.1109/TNSRE.2023.3324252
Problem

Classifying hand movements based on sparse multichannel sEMG signals remains challenging, despite recent advancements in deep learning techniques. Existing approaches typically involve single-model architectures, limiting their ability to extract representative features effectively. Addressing this limitation is crucial for developing accurate HGR systems, particularly for applications like myoelectric-controlled prostheses. Objective The objective of this study is to enhance hand gesture recognition (HGR) accuracy using surface electromyogram (sEMG) signals through a hybrid deep learning framework based on transformer architecture, called TraHGR. This framework aims to improve upon existing deep learning techniques for HGR classification, which often struggle with sparse multichannel sEMG signals. Methodology Framework Design: The proposed TraHGR architecture comprises two parallel paths followed by a fusion center consisting of a linear layer. This hybrid approach combines the advantages of both paths to enhance feature extraction and improve classification accuracy. Dataset: Evaluation is performed on the second Ninapro dataset (DB2), which contains sEMG signals from 40 healthy users performing 49 gestures under real-life conditions. Experimental Setup: Extensive experiments are conducted to validate TraHGR's performance and compare it with other HGR classification algorithms. Recognition accuracies are measured using a window size of 200ms and a step size of 100ms.Performance Evaluation: TraHGR's performance is compared with state-of-the-art methods across different subsets of the DB2 dataset. Additionally, the effectiveness of transformers for sEMG-based HGR is explored through comprehensive comparisons with traditional machine learning (ML) and deep neural network (DNN) approaches. Finding Performance Improvement: TraHGR achieves recognition accuracies of 86.00%, 88.72%, 81.27%, and 93.74% for different subsets of the DB2 dataset, outperforming state-of-the-art methods by significant margins. Hybrid Architecture Advantage: The hybrid architecture of TraHGR demonstrates superior feature extraction capabilities compared to single-path models, highlighting the effectiveness of combining multiple approaches for HGR.Transformer Effectiveness: The study illustrates the potential of transformer-based architectures for sEMG-based HGR, leveraging their success in other domains like natural language processing (NLP), computer vision (CV), and speech recognition.

2.5. FMCW Radar-Based Real-Time Hand Gesture Recognition System Capable of Out-of-Distribution Detection

Author: Jae-Woo Choi; Chan-Woo Park;
Year:2022
Doi: 10.1109/ACCESS.2022.3200757
Problem

Existing radar-based hand gesture recognition systems prioritize high classification accuracy but often lack the ability to detect OOD samples. In real-world HCI scenarios, misclassifying unintended gestures can lead to errors, particularly in real-time situations. Therefore, there is a need for systems capable of both accurate classification and OOD detection to ensure reliability. Objective This study aims to develop a real-time hand gesture recognition system using Frequency Modulated Continuous Wave (FMCW) radar sensors. The system not only classifies hand gestures accurately but also detects out-of-distribution (OOD) samples, enhancing its reliability for human-computer interaction (HCI) applications. Methodology Radar Data Processing and Classification: The study designs a radar data processing technique and employs a Transformer encoder-based classifier to achieve high classification accuracy for hand gestures detected by the FMCW radar sensor.OOD Detection: A relative Mahalanobis distance (RMD)-based OOD detection method is integrated into the system to identify OOD samples and enhance reliability. Dataset Collection: Three datasets are collected: an in-distribution dataset (Coarse) and two OOD datasets (Fine and Sign) to verify the proposed system's performance. Experimental Evaluation: Classification experiments are conducted using various network types and input types (image or sequence) on the Coarse dataset. OOD detection experiments are performed with different combinations of classifiers and OOD detection methods on the Coarse dataset versus the Fine and Sign datasets. The system's feasibility is demonstrated through real-time experimental demonstrations. Finding Classification Accuracy: The proposed system achieves a classification accuracy of 93.95% on the in-distribution dataset (Coarse). OOD Detection Performance: OOD detection experiments report AUROC values of 92.96% and 92.84% for the Fine and Sign datasets, respectively, indicating robust OOD detection capabilities. Feasibility: Real-time experimental demonstrations confirm the system's effectiveness in classifying complex hand gestures and detecting OOD samples, enhancing model reliability for real-world HCI scenarios.

2.6. Hardware Architecture Design for Hand Gesture Recognition System on FPGA

Author: Yuan-Chen; Po-Ting Chi
Year:2023
Doi:10.1109/ACCESS.2023.3277857
Problem

Effective HGR systems are crucial for touchless human-computer interaction. However, achieving high accuracy and fast computation with minimal hardware resources remains a challenge. This work addresses this issue by proposing a small-footprint HGR model and hardware architecture design. Objective This work aims to develop a compact and efficient hand gesture recognition (HGR) system for human-computer interaction, focusing on high recognition accuracy and computational speed. Methodology Model Design: The proposed HGR-Lite model integrates hand segmentation and gesture recognition tasks into a single-stage model, leveraging depth wise separable convolution to reduce parameters and computations. Hand segmentation is used as an attention mechanism to enhance recognition performance. Hardware Architecture: The hardware architecture of the neural model, including depth wise convolution, pointwise convolution, batch normalization, and max-pooling, is designed for efficient implementation on the Xilinx ZCU106 evaluation board. Implementation: The entire system is implemented on the evaluation board, achieving a performance of 52.6 frames per second (fps) and 65.6 giga operations per second (GOPS). Finding Model Performance: The HGR-Lite model achieves an accuracy rate of 89.25% on the OUHANDS test dataset, demonstrating its effectiveness in hand gesture recognition. Efficiency: By incorporating the segmentation model's features into the recognition model, the proposed approach significantly reduces parameters and computations while maintaining accuracy, making it suitable for resource-constrained environments.

2.7. Dynamic Hand Gesture Recognition Using Multi-Branch Attention Based Graph and General Deep Learning Model

Author: Abu Saleh Musa Miah; Md. Al Mehedi Hasan
Year:2022
Doi: 10.1109/ACCESS.2023.3235368
Problem

While previous approaches have utilized discriminative spatial-temporal attention features for hand gesture recognition from skeleton data, they often face challenges in achieving both high performance and generalizability due to inefficient feature extraction. Overcoming these challenges requires a novel approach capable of capturing diverse features effectively. Objective This study focuses on developing an efficient and high-performance hand gesture recognition system using dynamic hand skeleton data, which includes 3D coordinates of hand joints. The goal is to overcome the limitations of existing methods by proposing a multi-branch attention-based graph and a general deep learning model to extract diverse skeleton-based features for accurate gesture recognition. Methodology Model Architecture: The proposed model employs a multi-branch architecture consisting of two graph-based neural network channels and one general deep neural network channel. Each graph-based channel utilizes spatial and temporal attention modules to extract spatial-temporal and temporal-spatial features, respectively. The general channel extracts features using a standard deep neural network module. Feature Fusion: The spatial-temporal, temporal-spatial, and general features extracted from the branches are concatenated to form the final feature vector. Position embedding and mask operations are employed in the attention modules to track node sequences and reduce computational complexity. Evaluation: The proposed method is evaluated using three benchmark datasets: MSRA, DHG, and SHREC’17. Accuracy metrics are used to assess the performance of the model. Finding Performance Evaluation: The proposed approach achieves high accuracy rates of 94.12%, 92.00%, and 97.01% on the MSRA, DHG, and SHREC’17 datasets, respectively. These results demonstrate the effectiveness of the multi-branch attention-based graph and general deep learning model in capturing diverse features for gesture recognition. Comparison with State-of-the-Art: The experimental results indicate that the proposed method outperforms existing state-of-the-art methods in terms of both accuracy and computational cost.

2.8. Hand Gestures Recognition for Human-Machine Interfaces: A Low-Power Bio-Inspired Armband

Author: Andrea Mongardi; Fabio Rossi
Year:2022
Doi: 10.1109/TBCAS.2022.3211424
Problem

The problem addressed in this study revolves around the increasing demand for effective hand gesture recognition systems, particularly in the biomedical field where such systems serve as crucial components of Human-Machine Interfaces (HMIs). Despite the growing interest and development of wearable devices for gesture recognition, there exist challenges that hinder their widespread adoption and effectiveness. Objective This study aims to introduce a novel 7-channel surface Electromyography (sEMG) armband designed for hand gesture recognition, particularly targeting applications in serious gaming control and rehabilitation support. The focus is on achieving real-time gesture recognition with low power consumption and reduced complexity, using the Average Threshold Crossing (ATC) parameter as a bio-inspired technique. Methodology Armband Design: The prototype armband is designed to compute the Average Threshold Crossing (ATC) parameter, which counts how many times the sEMG signal crosses a threshold during a fixed time duration (130 ms) directly on the wearable device. This event-driven characteristic enables on-board prediction of common hand gestures with reduced power consumption. Functionality: The paper describes the physical structure, hardware/software components, and communication protocol of the armband. It also outlines the properties of the embedded machine learning algorithm, specifically an Artificial Neural Network (ANN), used for gesture recognition. Evaluation: The armband is tested with multiple configurations of the ANN, with the most efficient configuration achieving a testing accuracy of 91.9%. Latency due to prediction is measured at 1.34 ms, demonstrating suitability for real-time applications. Additionally, the armband's low current consumption of 2.92 mA allows for an operating time of up to 60 hours. Finding Gesture Recognition Accuracy: The proposed armband achieves an average classifier accuracy of 91.9% for recognizing eight active hand gestures plus an idle state, demonstrating competitive performance compared to existing solutions. Latency and Power Consumption: With a prediction latency of 1.34 ms, well below the 130 ms window of the ATC technique, and a low current consumption of 2.92 mA, the armband proves suitable for real-time applications with extended operating times.

III. PROPOSED SYSTEM :

The proposed system aims to revolutionize the interaction with PowerPoint presentations by leveraging hand gesture recognition technology. It provides a seamless and intuitive interface for controlling presentations through natural hand movements, offering increased mobility, responsiveness, and interactivity. Below are the key components of the proposed system: Gesture Net Model Integration The proposed system integrates the Gesture Net Model, a deep learning architecture specifically designed for hand gesture recognition. Trained on a dataset of hand gesture images or video clips, the Gesture Net Model offers high accuracy and robustness in classifying gestures. This integration enhances the system's gesture recognition capabilities, enabling precise and reliable interpretation of hand movements for controlling PowerPoint presentations. Two-Stream Network Architecture The system incorporates a Two-Stream Convolutional Neural Network (CNN) architecture for comprehensive gesture analysis. The Two-Stream Network consists of spatial and temporal streams, capturing both static spatial features and dynamic motion patterns of hand gestures. By leveraging this architecture, the system can effectively interpret both the static and dynamic aspects of hand movements, ensuring accurate recognition and classification of gestures in real-time. Integration with PowerPoint Software Seamless integration with popular presentation software such as Microsoft PowerPoint is a key feature of the proposed system. This integration allows direct control of slide navigation, start/stop functionalities, and other presentation actions through recognized hand gestures. By providing a plug-and-play solution, the system ensures compatibility with existing presentation setups without the need for extensive configuration, enhancing user convenience and accessibility. User Interface The proposed system offers an intuitive and user-friendly interface for configuring gesture controls and customizing mappings between gestures and actions. Users can easily calibrate and fine-tune the system according to their preferences, optimizing accuracy and responsiveness. This user-centric design approach ensures a seamless and personalized experience for presenters, facilitating smooth and efficient interaction with PowerPoint presentations.

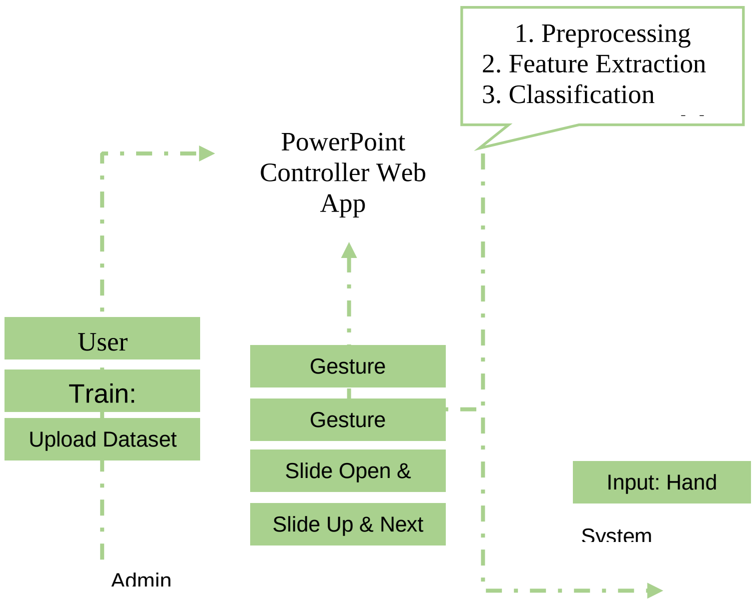


Figure 1: System Architecture of the proposed system

3.1 IMPLEMENTATION

Our project constituted of the below modules,

- Setting Up Development Environment
- Database Design and Setup
- Web App Development
- GestureNet Model Implementation
- Hand Gesture Recognition Setup
- Testing and Debugging
- Performance Evaluation
- Deployment

1.Setting Up Development Environment

Install necessary software and libraries such as Python, Flask, MySQL, Wampserver, TensorFlow, Pandas, Scikit Learn, Matplotlib, NumPy, Seaborn, Pillow, OpenCV, and Bootstrap. Configure development environment and ensure compatibility between different components.

2.Database Design and Setup

Design the database schema using MySQL to store user information, presentation data, gesture datasets, and model parameters. Set up the MySQL database and establish connections within the Flask application.

3.Web App Development

Develop the front-end interface of the PowerPoint Controller Web App using HTML, CSS, and JavaScript, with Bootstrap for responsive design. Implement user authentication and session management functionalities for secure access to admin and user dashboards. Design intuitive user interfaces for admin and user dashboards, incorporating features such as login, registration, presentation management, and gesture control.

4.GestureNet Model Implementation

Develop scripts for importing gesture datasets and preprocessing image data using OpenCV and NumPy. Implement feature extraction algorithms using convolutional layers, activation layers, and pooling layers with TensorFlow. Design and train the GestureNet Model using CNN architecture, optimizing model parameters using gradient descent algorithms. Integrate model deployment logic into the Flask application for real-time gesture recognition during presentation control.

5.Hand Gesture Recognition Setup

Configure video capture functionality using OpenCV to capture live hand video from the webcam or input device. Implement preprocessing steps to resize video frames, convert to grayscale, and apply noise reduction techniques. Develop algorithms for hand gesture recognition using Two Stream Networks with the trained GestureNet Model. Integrate gesture recognition logic with the PowerPoint Controller Web App for seamless control of PowerPoint presentations based on recognized gestures.

6. Testing and Debugging

Conduct thorough testing of each module and component to ensure functionality, accuracy, and reliability. Debug any issues or errors encountered during testing and address them promptly to improve system stability.

7. Performance Evaluation

Evaluate the performance of the hand gesture recognition system using metrics such as accuracy, precision, recall, and F1-score. Analyze the results of performance evaluation to identify areas for improvement and optimization.

8.Deployment

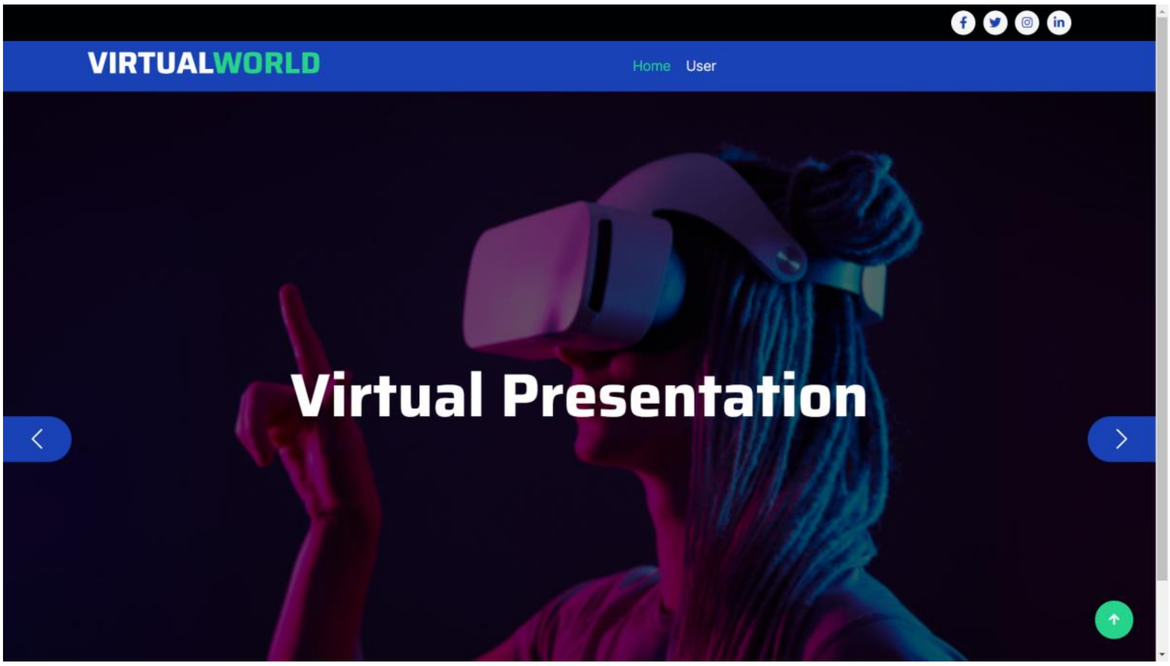
Deploy the fully developed and tested PowerPoint Controller Web App on a production server or cloud platform. Ensure proper configuration and scalability to handle user traffic and data storage requirements effectively. Monitor system performance post-deployment and address any issues that arise to maintain optimal functionality**.**

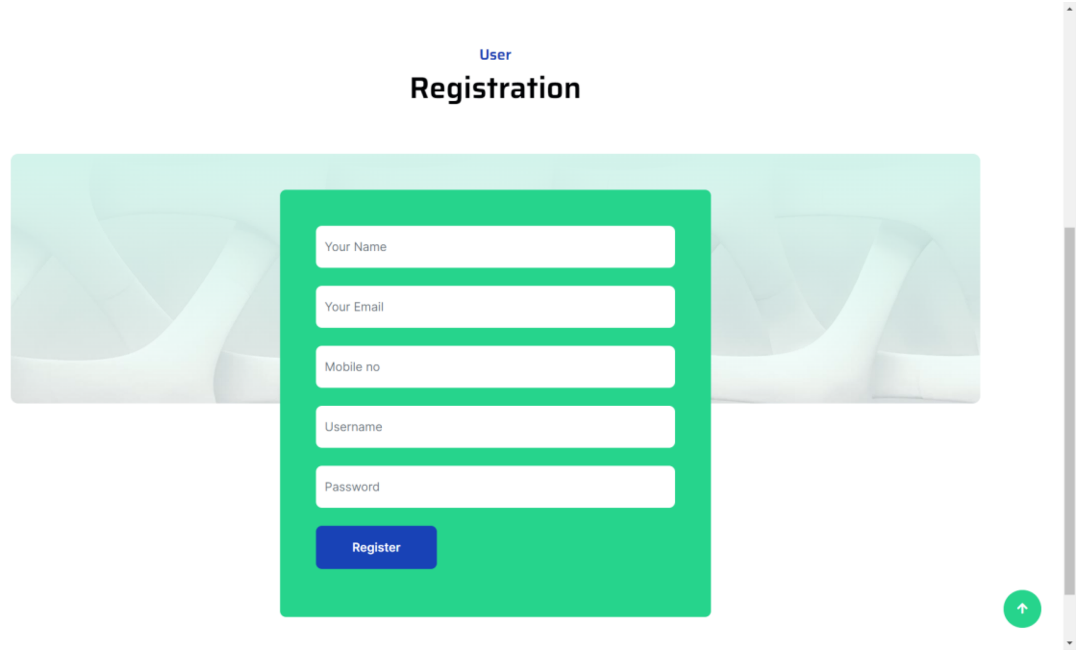
RESULTS AND DISCUSSION :

The purpose of testing is to discover errors and ensure that software systems meet their requirements and user expectations without failing in an unacceptable manner. Testing aims to uncover every conceivable fault or weakness in a work product, providing a means to check the functionality of components, sub-assemblies, assemblies, and finished products. Various types of tests address specific testing requirements, including unit testing, functional testing, acceptance testing, and integration testing. Unit testing focuses on validating internal program logic and ensuring that program inputs produce valid outputs, while functional testing systematically demonstrates that functions are available as specified by business and technical requirements. User acceptance testing (UAT) is critical for confirming that the system meets functional requirements, requiring significant end-user participation. Integration testing ensures that different software components or modules interact correctly and that the system as a whole meets functional and non-functional requirements. By combining these testing strategies, software testing provides comprehensive validation that each unique path of a business process performs accurately, identified inputs and outputs are handled correctly, and interfacing systems or procedures function as expected.

CONCLUSION :

The project successfully developed a comprehensive solution for controlling PowerPoint presentations using hand gestures, integrating web development, deep learning, and real-time video processing to enhance user interaction. Feasibility analysis and rigorous testing confirmed the system's functionality, reliability, and performance, providing an intuitive and interactive presentation experience.





REFERENCE :

1. K. S. Yadav, A. Monsley, K. Monsley and R. H. Laskar, "Gesture objects detection and tracking for virtual text entry keyboard interface", Multimedia Tools Appl., vol. 82, no. 4, pp. 5317-5342, Feb. 2023.
2. J. Gangrade and J. Bharti, "Vision-based hand gesture recognition for Indian sign language using convolution neural network", IETE J. Res., vol. 69, no. 2, pp. 723-732, Feb. 2023.
3. L. Liu, W. Xu, Y. Ni, Z. Xu, B. Cui, J. Liu, et al., "Stretchable neuromorphic transistor that combines multisensing and information processing for epidermal gesture recognition", ACS Nano, vol. 16, no. 2, pp. 2282-2291, Jan. 2022.
4. J. P. Sahoo, A. J. Prakash, P. Pławiak and S. Samantray, "Real-time hand gesture recognition using fine-tuned convolutional neural network", Sensors, vol. 22, no. 3, pp. 706, Jan. 2022.
5. Y. Jiang, L. Song, J. Zhang, Y. Song and M. Yan, "Multi-category gesture recognition modeling based on sEMG and IMU signals", Sensors, vol. 22, no. 15, pp. 5855, Aug. 2022.
6. Q. Gao, Y. Chen, Z. Ju and Y. Liang, "Dynamic hand gesture recognition based on 3D hand pose estimation for human–robot interaction", IEEE Sensors J., vol. 22, no. 18, pp. 17421-17430, Sep. 2022.
7. H. Ansar, A. Jalal, M. Gochoo and K. Kim, "Hand gesture recognition based on auto-landmark localization and reweighted genetic algorithm for healthcare muscle activities", Sustainability, vol. 13, no. 5, pp. 2961, Mar. 2021.
8. S. Saboo and J. Singha, "Vision based two-level hand tracking system for dynamic hand gestures in indoor environment", Multimedia Tools Appl., vol. 80, no. 13, pp. 20579-20598, May 2021.
9. P. Yu, J. Yin, Y. Sun, Z. Du and N. Cao, "An identity authentication method for ubiquitous electric power Internet of Things based on dynamic gesture recognition", Int. J. Sensor Netw., vol. 35, no. 1, pp. 57, Jan. 2021.

10. W. Juan, "Gesture recognition and information recommendation based on machine learning and virtual reality in distance education", J. Intell. Fuzzy Syst., vol. 40, no. 4, pp. 7509-7519, Apr. 2021.
11. Y. Wang, B. Zhang and C. Peng, "SRHandNet: Real-time 2D hand pose estimation with simultaneous region localization", IEEE Trans. Image Process., vol. 29, pp. 2977-2986, 2020.
12. Y. Li, X. Wang, W. Liu and B. Feng, "Pose anchor: A single-stage hand keypoint detection network", IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 7, pp. 2104-2113, Jul. 2020.
13. Y. Wang, C. Peng and Y. Liu, "Mask-pose cascaded CNN for 2D hand pose estimation from single color image", IEEE Trans. Circuits Syst. Video Technol., vol. 29, no. 11, pp. 3258-3268, Nov. 2019.
14. J. O. Pinzón-Arenas, R. Jiménez-Moreno and J. E. Herrera-Benavides, "Convolutional neural network for hand gesture recognition using 8 different EMG signals", Proc. 22nd Symp. Image Signal Process. Artif. Vis. (STSIVA), pp. 1-5, Apr. 2019.
15. S. Gattupalli, A. R. Babu, J. R. Brady, F. Makedon and V. Athitsos, "Towards deep learning based hand keypoints detection for rapid sequential movements from RGB images", Proc. 11th Pervasive Technol. Rel. Assistive Environments Conf., pp. 31-37, Jun. 2018.
16. P. Panteleris, I. Oikonomidis and A. Argyros, "Using a single RGB frame for real time 3D hand pose estimation in the wild", Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), pp. 436-445, 2018.
17. T. Simon, H. Joo, I. Matthews and Y. Sheikh, "Hand keypoint detection in single images using multiview bootstrapping", Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 4645-4653, Jul. 2017.
18. W. Wu, C. Li, Z. Cheng, X. Zhang and L. Jin, "YOLSE: Egocentric fingertip detection from single RGB images", Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW), pp. 623-630, Oct. 2017.
19. T. Grzejszczak, M. Kawulok and A. Galuszka, "Hand landmarks detection and localization in color images", Multimedia Tools Appl., vol. 75, no. 23, pp. 16363-16387, Dec. 2016.
20. Chatterjee, R. Xiao and C. Harrison, "Gaze+gesture: Expressive precise and targeted free-space interactions", Proc. ACM Int. Conf. Multimodal Interact., pp. 131-138, Nov. 2015.