PROPOSAL FORM
FINAL YEAR PROJECT IN COMPUTERSCIENCE/SOFTWARE ENGINEERING
Usman Institute of Technology University

AIR - TRACKER

HADIQA MEHDI	**21B-007-SE B**
MUSFIRA KHALID BUTT	**21B-027-SE B**
SAAD ATHER ALI	**21B-066-SE B**
DUA FAZIL	**21B-140-SE B**

SUPERVISED BY:

Department of computer Science

May 2024

PROPOSAL FORM
STUDENT INFORMATION FOR FINAL YEAR PROJECT

Fill in the information below:

Roll No	Enrollment No	Name	Email	Contact No
21B-007-SE	ST-21047	Hadiqa mehdi	21b-007-se@students.uit.edu	0311-3837022
21B-027-SE	ST-21107	Musfira Khalid butt	21b-027-se@students.uit.edu	0333-3517825
21B-066-SE	ST-21113	Saad Ather Ali	21b-066-se@students.uit.edu	0335-3948753
21B-140-SE	ST-21039	Dua Fazil	21b-140-se@students.uit.edu	0336-2223307

PROPOSAL FORM

Background	The proposed project focuses on developing a advanced user interaction tool for digital content as traditional presentation tools often rely on physical clickers or remote controls, which can be cumbersome and restrictive. Presenters often face challenges such as limited mobility, technical difficulties with remote controls, and the need for seamless integration of interactive features. These limitations can detract from the overall presentation experience affecting the productivity performance. **Air-tracker** addresses these challenges by leveraging advancements in AI and HCI to provide a more natural and engaging way to interact with presentation slides. By incorporating gesture recognition and real-time annotation capabilities, Air-tracker aims to enhance user experience and make presentations more dynamic and interactive. In the past, there have been several attempts to create more interactive presentation tools: **Leap Motion**: A motion control technology that allows users to interact with their computers using hand and finger motions. While innovative, it required a dedicated sensor and had limitations in gesture accuracy and user comfort. **Microsoft Kinect**: Originally designed for gaming, Kinect has been used in various applications, including presentations. However, it also required specific hardware and was more complex to set up. **Touchless Presentation Tools**: Various tools have used basic computer vision techniques to enable touchless interactions. These tools often lacked precision and required significant calibration and setup.
Project Scope	Air-tracker aims to develop an advanced presentation tool that redefines how users interact with digital slides. It will integrate innovative features to enhance engagement and interactivity during presentations. The scope of the Air-tracker project includes several key components: Data Preprocessing: Developing robust pipelines to clean and preprocess the captured video data for accurate gesture recognition. Gesture Recognition: Implementing advanced machine learning techniques to accurately detect and interpret hand gestures in real-time. Real-Time Annotation: Enabling real-time drawing and annotation on presentation slides based on recognized gestures. Accessibility: Air tracker enhances it for users with disabilities who may find it challenging to use traditional input devices. Intuitive User Experience: Air-tracker prioritizes user experience by offering an intuitive and natural way to interact with presentations.

Project Description	Air-tracker is a user-friendly presentation tool designed to elevate engagement and interactivity in presentations. It features intuitive gesture-based slide control, allowing users to navigate slides effortlessly through hand gestures instead of traditional clickers. Real-time drawing and annotation capabilities further enhance interaction by enabling users to annotate slides dynamically, fostering engagement and facilitating on-the-fly content customization. Air-tracker ensures seamless compatibility with various presentation formats and environments, leveraging advanced technologies like OpenCV for real-time gesture recognition and TensorFlow for machine learning models. By integrating these features, Air-tracker aims to redefine presentation experiences with a focus on usability, flexibility, and enhanced user interaction. Presenters need a more seamless, interactive, and engaging way to control their slides and annotate content in real-time, without being tethered to a podium or computer. There is a clear need for an innovative solution that leverages modern technologies. The main purpose of creating Air-tracker is to revolutionize the traditional presentation experience by introducing innovative, intuitive, and interactive methods of slide navigation and annotation. The tool aims to enhance user engagement and accessibility by replacing conventional clickers with gesture-based controls and enabling real-time drawing and annotation directly on slides. By leveraging advanced technologies like OpenCV, TensorFlow, and Media Pipe, Air-tracker provides a seamless and immersive presentation experience, making it easier for presenters to interact with their audience and convey their message more effectively.
Expected Technology	Hardware: Computer system(laptop), Camera. Software(s): Python for backend development Libraries: OpenCV, TensorFlow, Media Pipe for accurate hand tracking and gesture detection, NumPy. Algorithms: CNN, LSTM (future iterations)
Expected Outcome	The expected outcomes of Air-tracker include: A functional prototype that allows gesture-based control of presentation slides, enabling presenters to switch slides using hand gestures. A system capable of real-time drawing and annotation on slides, allowing presenters to make spontaneous annotations during their presentations. Enhanced user experience in presentations through intuitive and interactive controls, resulting in more engaging and effective presentations. Integration of OpenCV, TensorFlow, and Media Pipe for robust gesture recognition. Support for various presentation formats and environments, ensuring reliable performance.

Method/Approach	By following these steps, you can effectively build Air-tracker, ensuring it enhances user experience and accessibility during presentations through innovative gesture-based interactions. Data collection: Gather a diverse dataset of hand gestures and video annotations using tools like Labeling or VGG Image Annotator (VIA) for accurate labeling. Include gestures for slide control (e.g., swipe left, swipe right, stop) and drawing actions (e.g., point, draw). Data Preprocessing: Clean and preprocess the collected data to ensure high-quality inputs for training the machine learning models. Standardize input sizes, normalize pixel values, and apply data augmentation techniques like rotation, flipping, and scaling to enhance model robustness. Model Training: Utilize TensorFlow to train Convolutional Neural Network (CNN) models for gesture recognition. Leverage Media Pipe for hand tracking and detection, ensuring accurate real-time analysis of gestures. Integration: Develop the Air-tracker application using OpenCV for real-time image processing. Integrate this functionality seamlessly with popular presentation software to enable gesture-based slide control, real-time drawing, and annotation. Testing and Evaluation: Conduct extensive testing to validate the accuracy and reliability of gesture recognition and annotation features. Gather user feedback to refine and optimize the system based on usability and performance.
Relevant references	• Mittal, A., Sharma, S., & Tripathi, P. (2021). Hand Gesture Recognition for Human-Computer Interaction Using Computer Vision. ResearchGate. • Lee, J., & Kim, S. (2020). Gesture Recognition Techniques for Human-Computer Interaction: A Review," ProQuest. • Kumar, V., & Kumar, R. (2023). A Survey of Hand Gesture Recognition Techniques for Human-Computer Interaction. In *Advances in Computer Vision and Pattern Recognition*. Springer. • J. Lee et al., "Gesture recognition using OpenCV and TensorFlow," *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 567-589, 2023. • S. Brown, "Real-time annotation in digital presentations," *Journal of Human-Computer Interaction*, vol. 17, no. 3, pp. 345-359, 2022. • R. Clark, "Media Pipe: An open-source framework for multimodal machine learning," *IEEE Multimedia*, vol. 29, no. 2, pp. 45-55, 2022. • D. Davis, "Advancements in AI for presentation tools," *IEEE Transactions on Artificial Intelligence*, vol. 31, no. 1, pp. 23-35, 2023. • J. Anderson, "Advances in hand gesture recognition," in *Australasian Conference on Robotics and Automation*, 2009. • K. Smith et al., "Hand Gesture Recognition System Using Camera Libre," *International Journal of Computer Science and Information Security*, vol. 13, no. 8, pp. 25-30, 2015.

PROJECT DETAILS:

TECHNOLOGY PLATEFORM (TICK ONE OR MORE)
☒ Desktop application
☐ Web application
☐ Mobile application
☐ Gadget based Application
☐ Cloud based Application
☐Other : *Describe in your own context*

PROJECT STREAMS (TICK ONE OR MORE)
☐Block chain
☒ Image Processing
☒ Artificial Intelligence
☐ Computer Networks/wireless networks
☐Business application
☐CASE tools
☐System software
☐ Software Automation
☐Network Security
☐ AR/VR/3D Modeling
☐ Data Sciences/ Information retrieval and analysis
☐ Neural Network
☐ Others : ☒ HCI

Domains

☐ **Finance**
☐ **Ecommerce**
☐ **Banking**
☒ **Education**
☐ **Health**
☐ **Telecommunication**
☐ **Construction**
☐ **Insurance**
☐ **Other**

Endorsed By:
Their Suggestion and improvements:

Signature of faculty

Endorsed By:
Their Suggestion and improvements:

Signature of faculty

For Office use only (To be filled by the Final Year Project Coordinator)

Sr. No	Evaluator name	remarks	Accepted/ rejected

Idea Accepted	**Idea Rejected**
	Reason(s): Next Action Plan:

FYP COORDINATOR