

Engineering and Health Science Collaborative Model for Assessing Augmented Reality and Vision-Based Hand-Tracking Systems

Gayathri Boopathy, University of Massachusetts Lowell

Gayathri Boopathy is a Ph.D. student in Computer Engineering whose research focuses on augmented reality (AR), machine learning, physical therapy, MediaPipe, and human–computer interaction. Her work integrates immersive technology with Machine Learning and AI to design intelligent systems for motion analysis, gesture recognition, and rehabilitation support. Building on her previous research on AR-based hand tracking and movement classification, she explores the fusion of spatial computing and biomechanical modeling to advance digital health and interactive computing. Gayathri’s interdisciplinary approach bridges engineering, healthcare, and data science to develop impactful solutions for next-generation human–machine interfaces.

Emi Aoki, University of Massachusetts Lowell

Emi Aoki is a PhD student in Electrical Engineering at the University of Massachusetts Lowell. Her previous research experience has been in data analysis, computational modeling, and augmented reality application development and design. Her current research interests focus on modeling and simulation for digital twins and Model-Based Systems Engineering.

Flore Stécie Norcéide, University of Massachusetts Lowell

Flore Stécie Norceide is a PhD student in Electrical Engineering at the University of Massachusetts Lowell. Her research focuses on benchmarking hardware and software architectures for event driven computation with a focus on its application to object detection, tracking and pose estimation.

Ravi Kkumaar Sivasankar Venkata, University of Massachusetts Lowell

Ravi Kkumaar Sivasankar Venkata is a Master’s student in Computer Science at the University of Massachusetts Lowell. My work focuses on augmented reality, virtual reality, spatial computing, and human-computer interaction using AR devices and Unity. His research explores gesture-driven interaction systems, real-time hand tracking, immersive XR application and AI-integrated AR applications, with focus on rehabilitation, education and assistive technologies.

Dr. Erika S. Lewis, University of Massachusetts Lowell

Erika S. Lewis, PT, Ed.D., MS, CHT is an Associate Professor of Physical Therapy and Kinesiology Department at the University of Massachusetts Lowell. She has expertise as a physical therapist (licensed since 1992) and a Certified Hand Therapist (credentialed since 1998). Dr. Lewis is co-founder of the UMOVE UMass Movement Research Center. Her research interests are musculoskeletal hand and upper extremity injuries, Augmented Reality for hand tracking and rehabilitation, Virtual Reality for Rehabilitation, specifically Virtual Reality Wheelchair Training, and 3D scanning and 3D printing wrist orthoses for people with chronic wrist pain and Ehler’s Danlos Syndrome. Other research topics include treatment outcomes, orthosis fabrication, emotional intelligence, stenosing tenosynovitis (trigger finger), and reliability in instruments. Dr. Lewis teaches the Musculoskeletal Lower Extremity and Upper Extremity Evaluation series lectures and labs in the Doctor of Physical Therapy (DPT) Program. Additionally, she is co-founder of the Exercise Science elective called “Adaptive Devices for Better Life” an interprofessional experiential undergraduate course with interdisciplinary teams (PTK, Art & Design, Mechanical Engineering) who meet community members and design and create devices to make participation and activities easier for people with alternate abilities. One student group, ZipperBuddy won first place in 2023, winning \$5,000 at the Kennedy College of Engineering Prototype Competition to develop their invention. ZipperBuddy obtained a patent and became a company. Dr. Lewis is currently practicing at Rehab Health 360, Acton, MA where she evaluates and treats Hand and Upper Extremity patients including orthopedic, plastic surgery and neurological conditions and diagnoses.

Prof. Charles Thompson PhD, University of Massachusetts Lowell

Charles Thompson is Professor of Electrical and Computer Engineering, Director of the Center for Advanced Computation and Telecommunications and formerly Associate to the Dean for Research and Graduate Study at the University of Massachusetts at Lowell. He received a B.S. in Electrical Engineering from New York University, a M.S. in Electrical Engineering from the Polytechnic Institute of New York, and a Ph.D. in Acoustics from the Massachusetts Institute of Technology.

Upon graduation he became an Assistant Professor of Engineering Science and Mechanics at the Virginia Polytechnic Institute and State University. In 1987 he joined the Department Electrical and Computer Engineering at UMASS Lowell as its Analog Devices Career Development Professor. Dr. Thompson has served on the executive boards of the Cooperative Research Fellowship program of Bell Laboratories (1991-1999) and the AT&T Labs Fellowship Program (1996-2006).

At Bell Laboratories Dr. Thompson created with the Vice President of Research and Nobel laureate, Arno Penizas, the W. Lincoln Hawkins Mentoring Excellence Award (1994). This award is given to a member of the research staff for fostering the career growth of Bell Labs students and associates. This award is Research's highest honor for mentoring contributions. In 1998, AT&T Labs instituted a similar award named for Dr. Thompson.

Charles Thompson is Professor of Electrical and Computer Engineering, Director of the Center for Advanced Computation and Telecommunications and formerly Associate to the Dean for Research and Graduate Study at the University of Massachusetts at Lowell. He received a B.S. in Electrical Engineering from New York University, a M.S. in Electrical Engineering from the Polytechnic Institute of New York, and a Ph.D. in Acoustics from the Massachusetts Institute of Technology.

Upon graduation he became an Assistant Professor of Engineering Science and Mechanics at the Virginia Polytechnic Institute and State University. In 1987 he joined the Department Electrical and Computer Engineering at UMASS Lowell as its Analog Devices Career Development Professor. Dr. Thompson has served on the executive boards of the Cooperative Research Fellowship program of Bell Laboratories (1991-1999) and the AT&T Labs Fellowship Program (1996-2006).

At Bell Laboratories Dr. Thompson created with the Vice President of Research and Nobel laureate, Arno Penizas, the W. Lincoln Hawkins Mentoring Excellence Award (1994). This award is given to a member of the research staff for fostering the career growth of Bell Labs students and associates. This award is Research's highest honor for mentoring contributions. In 1998, AT&T Labs instituted a similar award named for Dr. Thompson.

His awards include the US Presidential Award for Excellence in Mentoring; Tau Beta Pi Eminent Engineer; James E. Blackwell Scholar; AT&T Bell Laboratories Cooperative Research Fellowship. He is cited in Who's Who among African Americans, Education, and Technology Today; American Men and Women of Science, West Babylon Alumni Hall of Fame; He is a Fellow of the Acoustical Society of America and cited for his fundamental contributions to theoretical and computational acoustics. He is senior member of IEEE, and a member of the American Physical Society and Sigma Xi. He has published research in acoustics, control theory, fluid mechanics, heat transfer, linear and nonlinear systems, and telecommunications.

Prof. Kavitha Chandra, University of Massachusetts Lowell

Kavitha Chandra is the Associate Dean for Undergraduate Affairs and Professor of Electrical and Computer Engineering in the Francis College of Engineering at the University of Massachusetts Lowell. She directs the Research, Academics and Mentoring Pathways (RAMP) to Success summer bridge and academic program for new engineering students, preparing them with research, communication and leadership skills. Her research interests are in computational and data-driven modeling of physical systems in acoustics and communication networks, model-based systems engineering, user-centric design of emerging technology and engineering education.

Engineering and Health Science Collaborative Model for Assessing Augmented Reality and Vision-Based Hand-Tracking Systems

Abstract

The potential of digital health solutions with immersive technology that can respond to the growing demand for remote healthcare, such as physical therapy, motivates the need for educational frameworks that prepare engineering students to be better informed about designing healthcare applications. The focus is on Augmented reality (AR) and computer-vision-based hand tracking systems, both of which offer promising solutions for remote rehabilitation by enabling real-time capture of hand motion and gesture data; however, they differ in sensing modality, accuracy, latency, cost, and implementation complexity. This work presents an interdisciplinary educational framework that integrates clinically relevant information from health sciences to inform the design and assessment of technology based applications for hand tracking. The modules are designed to be accessible to undergraduate students in engineering and health-sciences, promoting increased collaboration between these disciplines. The modular framework enables students to implement, analyze, and compare 3D sensor-fusion-based AR tracking with 2D computer-vision-based tracking, focusing on clinically relevant metrics such as joint angles, gesture classification, and range-of-motion estimation. Two rehabilitation-inspired, project-driven activities—gesture-based object manipulation in a spatial grid and robotic arm control—provide hands-on experiential learning aligned with physical therapy concepts and human-computer interaction principles. The educational content is derived from ongoing joint research by graduate students in engineering and physical therapy, with licensed physical therapists guiding the process. By embedding immersive technologies within a structured, outcome-based instructional model that is also open-access, this work demonstrates a scalable approach to teaching the design of digital health technologies that incorporate engineering, rehabilitation, and human-centered design.

I. Introduction

The growth in the aging population, the increased prevalence of musculoskeletal and neurological conditions, and persistent barriers to in-person care, such as cost, geography, and availability of clinicians, have created unprecedented demand for accessible and scalable physical therapy (PT) solutions [1, 2]. The potential of digital health technologies that support remote PT can be significant in addressing these barriers. However, the time required for the adoption of technological solutions in health-care can take decades [3, 4]. This process may be accelerated through educational systems that promote interdisciplinary collaboration between technology designers and health science experts.

This work presents a collaborative model between researchers in engineering and health sciences that has been implemented at the University of Massachusetts Lowell. The research problem addressed is the design and evaluation of multi-sensor based augmented reality (AR)-based applications as digital health solutions for remote PT to rehabilitate hand mobility. The AR application is compared with a software based computer vision application designed with Google's MediaPipe application [5].

AR head-mounted devices integrate multiple sensors to estimate the three-dimensional pose of hands in real time, allowing for spatial interaction and joint-level motion tracking within an immersive environment [6]. In contrast, vision-based algorithmic approaches use monocular RGB cameras and machine learning models to infer hand landmarks from two-dimensional video streams, offering a low-cost and widely accessible alternative for motion analysis [7, 8].

Although both approaches have demonstrated promise, they differ substantially in the sensing modality,

accuracy, latency, depth estimation, and computational hardware that directly influence their suitability for clinical decision-making and rehabilitation monitoring [1]. To better understand the relevant clinical decision variables, a research collaboration between engineering and health science researchers with expertise in PT was initiated. This paper describes the progression of this multidisciplinary project and the educational modules that were designed from this collaboration. The sensor-fusion-based AR and computer-vision-based MediaPipe using webcam system evaluation research will be referred to as the AR-CV project.

The two assessments conducted involve: (i) three-dimensional hand tracking using the Magic Leap 2 AR headset integrated with Unity, a 3D real-time development platform, and (ii) two- and three-dimensional hand tracking using MediaPipe on a flat-screen interface. For each system, the hand mobility is recorded through observations of the distal interphalangeal (DIP), proximal interphalangeal (PIP), interphalangeal (IP), and metacarpophalangeal (MCP) joints of the fingers during prescribed therapeutic movements and analyzed over time.

Previous studies on engineering-health science collaboration identify both enabling factors and challenges, including communication barriers, differing validation standards, and the need for mutual trust and shared objectives [9, 10, 11]. Successful models emphasize the early involvement of clinicians, iterative feedback, and the integration of domain expertise throughout the research lifecycle.

To develop viable solutions to the AR-CV problem, engineering researchers acquired the necessary background in hand physiology, types of mobility impairments, and related exercise regimens proposed by physical therapists. The confidence of clinical science researchers in technology-based solutions from engineers was critical. This was accomplished over time through sustained communication and demonstrations of technology capabilities across the design cycle, generating clinically applicable results.

Fig. 1 shows the sequence of events that took place over the last two-year period. The collaboration originated through the development of joint proposals between engineering and health science faculty, followed by graduate student research focused on AR-enabled hand tracking and clinician-guided study design. Licensed physical therapists and PT students actively contributed to the definition of clinically meaningful joint metrics that were applied to validate AR and MediaPipe-based measurements against standard goniometric assessments. Human-subject studies were conducted to collect comparative data across the AR and MediaPipe platforms, supporting an iterative process of technical refinement informed by clinical insight.

An example of the collaboration between engineering and PT students included the demonstration of

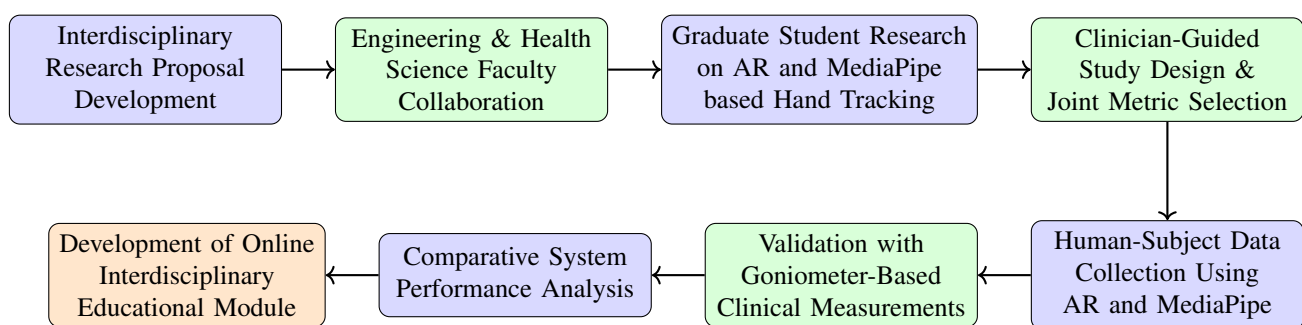


Figure 1. Flow diagram illustrating the interdisciplinary collaboration model that integrates engineering research, clinical validation, and educational deployment. (Engineering activities are shown in blue, clinical activities in green, and educational outcomes in orange.)

the application of the goniometer that PTs utilize to measure the mobility of the finger joints. Fig. 2 shows a panel that depicts the goniometer measurement process carried out by a graduate student in PT, recording the finger-joint angle made by the four fingers of a volunteer during the human subject studies. The measurements were recorded by the engineering student who also designed the AR and MediaPipe experiments in which the volunteers participated.

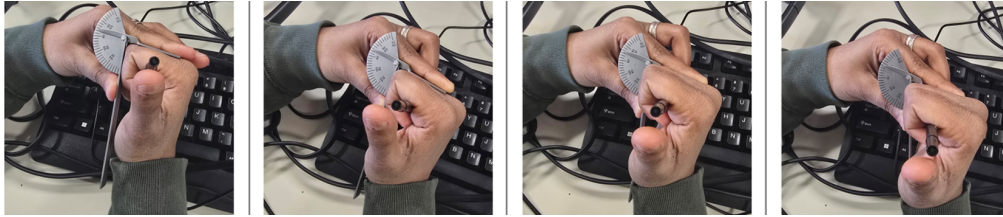


Figure 2. Physical therapist making measurements of the proximal interphalangeal (PIP) joint angles using a goniometer for a thumbs-up right-hand gesture. Conducted during the human-subject activity stage, the objective was to provide comparative measures against AR and MediaPipe generated measurements.

The research process and results derived are mapped to online educational modules on the Canvas Platform. They describe the clinical and engineering concepts in the proposed design. From a more general educational perspective, AR and gesture-based interaction have been shown to improve learning, engagement, and skill acquisition in engineering, medical, and rehabilitation education [6, 12, 13]. By embedding rehabilitation-inspired projects, the proposed modules support and extend experiential learning, building skills at the intersection of immersive systems, human-computer interaction, and digital health.

Section II provides a review of related work in hand tracking, immersive learning, and interdisciplinary digital health education, that informs the design of the proposed educational framework. Section III presents the educational modules for acquiring background knowledge on developing hand tracking applications with two types of technology. Section IV presents two interactive, project-driven activities that demonstrate gesture-based virtual object manipulation and robotic arm control using both AR and vision-based platforms, respectively. Section V concludes with a discussion on the interdisciplinary collaboration and educational deployment.

II. Background and related work in hand tracking and interdisciplinary engineering education

Zahid et al. [14] propose an immersive learning approach in which students learn through real-time kinesthetic hand interaction in extended reality (XR) environments rather than passive content consumption, demonstrating improved engagement and self-directed exploration. Eiler et al. [15] discuss pedagogical design and learning effectiveness when integrating AR and Virtual Reality (VR) technologies into engineering education courses.

In a engineering manufacturing laboratory context, Grodotzki et al. [6] demonstrate that human controller-free interaction with AR improves laboratory learning, spatial understanding, and usability through interactive overlays. Connolly et al. [16] show that mixed reality platforms such as HoloLens can deliver structured remote clinical instruction with measurable learning gains. Husna et al. [2] design gesture-controlled rehabilitation exercise games that improve engagement and support movement learning through repetition and feedback. Zechner et al. [17] further demonstrate that performance metrics and effort-related signals can be used as evidence when assessing learning and training in mixed-reality environments.

Zhang et al. [7] introduce an open, real-time hand skeleton tracking framework using 21 RGB input landmarks, which has since become a “hand-tracking engine” foundation for educational demonstrations and rehabilitation-oriented applications.

MediaPipe-based hand tracking has been shown to provide objective quantitative feedback in simulation-based medical training [12] and it has been used to design gesture-driven educational activities in which hands serve as the primary input modality for learning [13]. The educational impact of hand tracking was further expanded by Alsharif et al. [18] by enabling real-time interpretation of American Sign Language to improve accessibility and inclusion.

In rehabilitation and assessment contexts, Sasha et al. [8] report the use of MediaPipe Hands to estimate the joint range of motion of MCP, PIP, and DIP. Amprimo et al. [1] emphasizes that vision-based hand tracking must be carefully validated before being considered suitable for rehabilitation-grade measurement.

The pedagogical benefits and persistent challenges related to accessibility, scalability, and rigorous learning assessment with the use of AR and XR in education have been reviewed [19]. Finally, Lei et al. [20] demonstrates that web-based digital twin laboratories can support remote, interactive engineering education without specialized hardware.

From an interdisciplinary perspective, Boopathy et al. [21] proposed a digital engineering framework that integrates clinicians and system engineers to model, validate, and quantitatively assess immersive rehabilitation technologies using clinically meaningful metrics such as finger-joint angles [22]. Extending this interdisciplinary philosophy into education, the present work embeds clinician-defined metrics directly into student-facing modules, allowing learners to experience how clinical requirements shape engineering design.

Collectively, the literature supports the use of immersive and vision-based hand tracking technologies to improve student engagement, interactive learning, and quantitative skill development across engineering, healthcare, and rehabilitation contexts. Previous studies also demonstrate the value of AR/VR environments, gesture-based interaction, and motion-analysis as effective tools for experimental education and training. At the same time, existing work provides limited information on the direct comparison between AR and vision-based systems, the integration of clinically relevant rehabilitation tasks, and interdisciplinary instructional models that bring together engineering, health science, and creative design perspectives within a single framework. These gaps motivate the AR-CV research, along with rehabilitation-inspired learning activities and open-access modules, to enable scalable interdisciplinary engineering education.

The present work adopts gesture-driven interaction as the overarching problem to be studied in the instructional approach and integrates clinically informed hand kinematics, joint-angle computation, and quantitative feedback into engineering activities. Identical tasks are designed for direct comparison between immersive and non-immersive interaction. Rehabilitation-inspired hand tasks motivate learning about sensor function, kinematics, and human-computer interaction, while introducing clinically meaningful performance metrics.

III. Educational modules integrating immersive systems applications for hand tracking

The modules are designed for collaboration between students in engineering, PT, and exercise disciplines. Engineering students develop an understanding of hand anatomy, joint kinematics, and interaction

Course Introduction

This course presents a comparative study of hand tracking using Augmented Reality (AR) platforms and MediaPipe.

Learners will explore the fundamentals of hand tracking, understand software and hardware requirements, and implement two real-world use cases using both technologies.

The course emphasizes hands-on implementation and performance comparison between AR-based and computer-vision-based approaches.

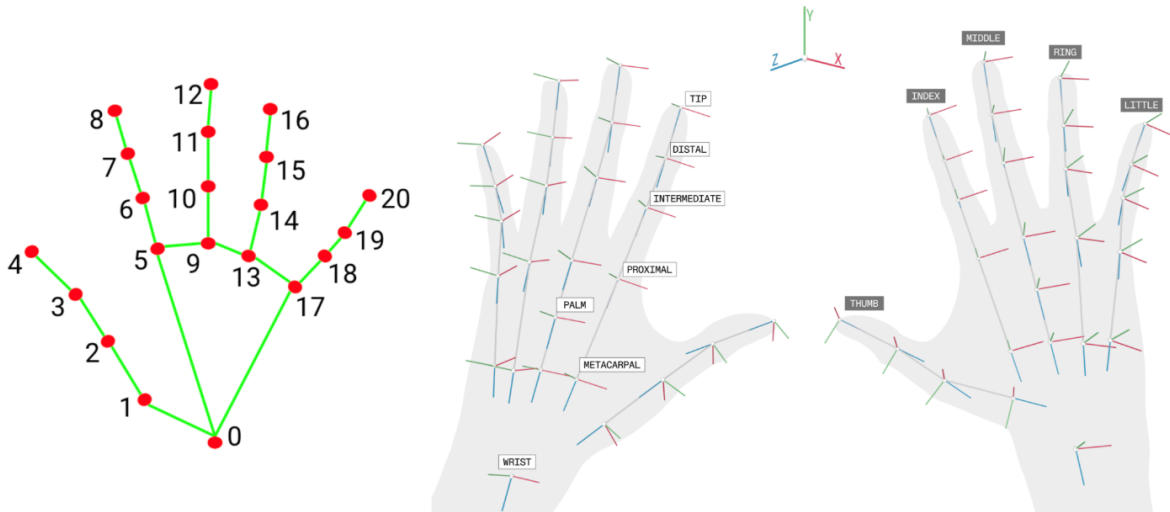


Figure 3. Course Introduction Page in Canvas Module

design principles to build functional and accurate hand-tracking tools. PT students gain exposure to computational methods and sensing technologies, leading to knowledge about how digital systems can support rehabilitation assessment and therapy design.

The instructional framework follows outcome-based education principles and aligns with Bloom's taxonomy [23] by gradually progressing from the acquisition of foundational knowledge to higher-order skills, including application, analysis, and evaluation. Emphasis is placed on experiential learning, allowing students to develop problem-solving abilities relevant to emerging fields such as extended reality and intelligent human-computer interaction.

The course is designed as a sequence of online learning modules on the public Canvas platform. This modular organization supports self-paced learning while ensuring logical continuity between conceptual topics, implementation exercises, and comparative analysis. The Canvas environment also facilitates continuous evaluation through quizzes, laboratory submissions, and a final project.

The course starts with a *“Getting to Know Participants/Students”* module to understand learners' backgrounds, prior experience, motivation, and expectations. This helps instructors adapt course content, activities, and technical depth to better support students' learning needs.

A. Framing the hand-tracking problem

The course introduction page shown in Fig. 3 frames the problem, depicting the hand joints of interest for tracking and analysis. The organization of the introductory modules is shown in Fig. 4. Module 1 introduces Hand Tracking, information on hand physiology emphasizing the finger joints, and a module on the technology requirements. The introductory module establishes the scope, tools, learning outcomes, and assessment methodology. This module familiarizes AR-CV project learners with the overall objectives of the course and the platforms used throughout the curriculum. Following this

- Module 0: Course Overview & Setup - Course Introduction - Learning Outcomes - Tools & Platforms Overview - Module 1: Fundamentals of Hand Tracking - What is Hand Tracking? - Hand Anatomy for Interaction Design - Joint Motion and Kinematic Analysis - Watch demo videos of hand tracking in AR and MediaPipe - Module 2: Core Principles of Hand Tracking and Gestures - Links/References - AR & Mediapipe - Hand Gesture Types (Static vs Dynamic) - Coordinate Systems & Depth Estimation	- Module 3: Software & Hardware Setup - Software and Hardware Module - Hardware Requirements - AR Devices - MediaPipe Hardware - Software Requirements - AR Platforms - MediaPipe Software - Environment Setup - Installing tools for AR - Installing tools for MediaPipe - Setup verification 100 pts
---	--

Figure 4. Canvas Course Introduction and Requirements/Setup Module

introduction, a foundational Module 2 introduces the core principles of hand tracking, including the distinction between sensor-fusion-based and computer-vision-based approaches. Concepts such as static and dynamic gestures, as shown in Fig. 5, coordinate systems, and depth estimation are discussed to provide learners with the theoretical grounding required for platform-specific development.

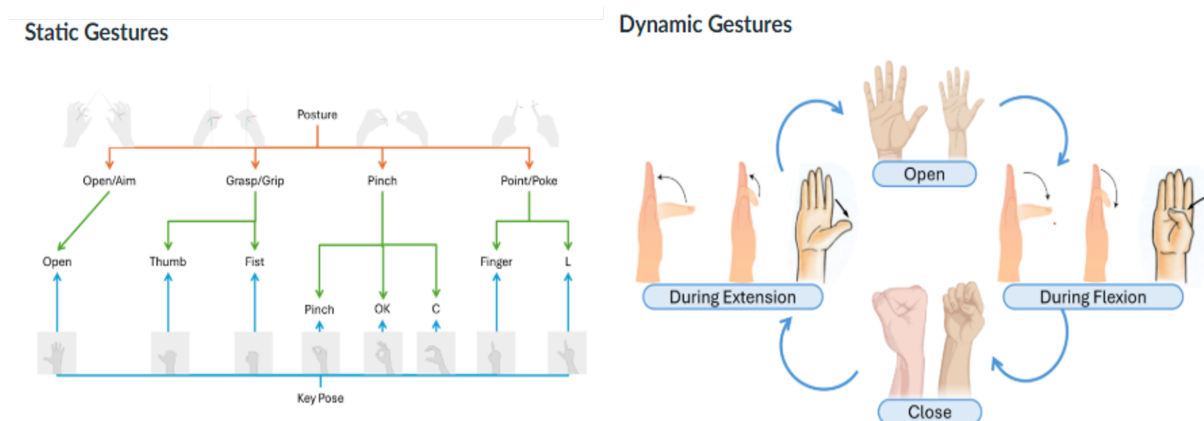


Figure 5. The left image illustrates static hand gestures organized hierarchically by posture and key poses, including open/aim, grasp/grip, pinch, and point/poke with their corresponding sub-gestures. The right image depicts dynamic gestures, showing temporal hand motion transitions between open and close states during extension and flexion, emphasizing gesture recognition over time as presented on the Canvas page.

The hardware and software setup/requirements for implementing hand tracking on both platforms are described in Module 3. This section provides a comparative discussion of the AR-CV systems. The module explains the sensing capabilities, development environments, and configuration processes associated with each platform. The learners are guided through the installation and setup procedures, ensuring readiness for subsequent laboratory activities. A setup verification task is included to confirm a successful environment configuration and to minimize technical barriers at later stages of the course.

B. AR-CV hand tracking modules

The course devotes two full modules to hand tracking implementation using Magic Leap 2 and MediaPipe as shown in Fig. 6. Module 4 enables the student to explore the sensor-fusion-based architecture of the AR device, explaining how multiple sensors contribute to accurate three-dimensional hand tracking. Learners examine hand landmark detection and gesture recognition mechanisms within a spatial computing environment. Practical laboratory exercises enable students to visualize hand joints and implement gesture-based interactions using Unity and the Magic Leap SDK. This module emphasizes real-world spatial interaction and highlights platform-specific constraints, such as occlusion and field-of-view limitations.

Module 4: Hand Tracking in AR Platforms	Module 5: Hand Tracking Using MediaPipe
AR Hand Tracking Architecture	MediaPipe Framework Overview
Hand Landmark Detection	Hand Landmark Model
Gesture Recognition in AR	Gesture Classification Logic
Limitations (occlusion, lighting, device dependency)	Performance and Accuracy Considerations
Detect hand landmarks in AR 50 pts	Hand landmark detection using webcam 50 pts
Visualize joints and fingers 50 pts	Gesture recognition using MediaPipe 50 pts

Figure 6. Canvas Hand Tracking Module (AR and MediaPipe)

Module 5 demonstrates hand tracking with the MediaPipe application and a standard webcam. It describes the MediaPipe framework and its processing pipeline, including palm detection, hand landmark estimation, and post-processing stages. Learners study the 21-keypoint hand landmark model and develop gesture recognition logic based on relative joint positions and motion patterns. Performance considerations such as lighting conditions, camera resolution, and computational limitations are discussed to highlight practical challenges in computer-vision-based hand tracking. A hands-on laboratory activity enables learners to implement real-time hand tracking and gesture detection using Python. The panel in Fig. 7 highlights some of the visualizations integrated within these modules.

C. Comparative evaluation and module learning outcomes

To consolidate learning, Table. I depicts a high level evaluation of AR-CV across multiple system dimensions, including accuracy, latency, hardware dependency, ease of development, scalability, and cost. This analytical subsection encourages students to critically assess the strengths and limitations of

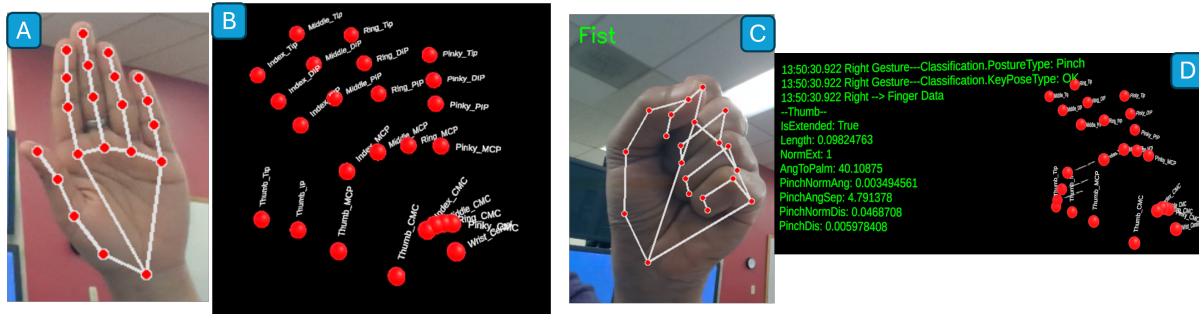


Figure 7. A comparative visualization of hand landmark tracking and gesture recognition using MediaPipe and an AR-based framework. Fig A shows an open hand detected by MediaPipe with 21 skeletal joints represented as red dots. Fig B illustrates an open hand in the AR system, tracking 26 hand joints and displaying their corresponding landmark names in a virtual view. Fig C demonstrates MediaPipe's gesture classification for a closed hand, identified as a Fist gesture. Fig D shows the AR-based gesture classification of an OK/Pinch gesture, along with real-time system logs in green text that record pinch-related parameters and index-thumb finger interaction data within the virtual environment.

Table I
OVERALL COMPARISON BETWEEN AR AND MEDIAPIPE

Platform	Tracking	Depth	Accuracy	Latency	Cost	Dev. Effort	Portability	Best Use Case
AR	Sensor fusion	Yes	Very high	Low	High	High	Limited	Spatial AR
MediaPipe	Computer vision	No	High	Medium	Very low	Low	High	Desktop / Web apps

each platform in the context of engineering applications. By directly comparing sensor-fusion-based and vision-based approaches, students develop the ability to make informed decisions regarding technology selection for specific use cases and deployment scenarios.

Table II presents the general structure of the online educational modules and their associated student-centered learning outcomes. The module sequence is designed to guide learners from the foundational concepts in hand tracking to platform-specific implementation, project-driven activities, and final comparative evaluation. By integrating AR-CV approaches within a unified instructional framework, the module supports experiential learning, interdisciplinary collaboration, and critical analysis of system-level tradeoffs relevant to engineering and rehabilitation applications.

The next section presents two educational project-driven activities that translate conceptual differences between AR and CV into hands-on, task-driven learning experiences that highlight practical implications for engineering design and health science applications.

IV. Project-driven activities

The project-driven activities are delivered through an online description and enable remote access interaction to both AR and MediaPipe hand tracking modules. MediaPipe activities can be completed fully using a standard webcam. AR-based activities are supported through a cloud-based interaction, where students use Unity-based XR simulators or AR device application simulators for development and testing. Applications can then be deployed and evaluated on AR devices through cloud-connected workflows, allowing instructors to monitor execution, verify functionality, and support student progress remotely.

To ground the comparative evaluation of AR and CV systems, two activities were developed. These projects are motivated by upper-extremity rehabilitation and PT practice, where an accurate assessment

Table II
STRUCTURE AND LEARNING OBJECTIVES OF THE ONLINE EDUCATIONAL MODULE

Module	Focus	Learning Outcome
Module 0: Course Overview & Setup	Course introduction, structure, tools, and interdisciplinary context	Students understand the course scope, learning objectives, assessment structure, and overview of AR and MediaPipe platforms.
Module 1: Fundamentals of Hand Tracking	Basic concepts of hand tracking, hand anatomy, and motion principles	Students understand hand tracking fundamentals, hand anatomy for interaction design, and joint motion relevant to gesture-based systems.
Module 2: Core Principles of Hand Tracking and Gestures	Gesture representation, coordinate systems, and depth estimation	Students understand gesture types (static vs dynamic), coordinate systems, and depth estimation techniques used in AR and vision-based tracking.
Module 3: Software & Hardware Setup	Comparative hardware and software setup for AR and MediaPipe	Students identify, compare, and configure development environments, hardware requirements, and software tools for AR-based and MediaPipe systems.
Module 4: Hand Tracking in AR Platforms	Sensor-fusion-based 3D hand tracking in AR	Students implement hand landmark detection and gesture recognition in AR systems and evaluate platform-specific constraints.
Module 5: Hand Tracking Using MediaPipe	Computer-vision-based hand tracking using RGB cameras	Students implement hand landmark detection and gesture classification using MediaPipe and evaluate performance in real-world conditions.
Module 6: Platform Comparison - AR vs MediaPipe	System-level comparison across platforms	Students analyze differences in accuracy, latency, portability, development effort, and cost between AR and MediaPipe approaches.
Module 7: Project 1 - Object Placement (AR)	Gesture-based object manipulation in a 3×3 spatial grid using AR	Students design and implement gesture-driven object interaction in a three-dimensional AR environment.
Module 8: Project 1 - Object Placement (MediaPipe)	Vision-based object manipulation in a 3×3 grid using MediaPipe	Students implement gesture-driven object placement using vision-based interaction and compare interaction fidelity.
Module 9: Project 2 - Robotic Arm Control (AR)	Gesture-controlled robotic arm using AR	Students map sensor-fusion-based hand gestures to robotic joint motion, implement control logic, and use MQTT for communication.
Module 10: Project 2 - Robotic Arm Control (MediaPipe)	Robotic arm control using MediaPipe hand tracking	Students implement hand gesture-based robotic control using vision-based tracking and evaluate smoothness and latency.
Module 11: Final Comparison and Evaluation	Cross-platform evaluation and final project	Students critically evaluate AR and MediaPipe implementations and present a final project with analysis and demonstration.

of hand and finger kinematics is essential. In health science and PT contexts, this approach provides a fast, digital alternative to traditional goniometer-based joint angle measurement for rehabilitation assessment and training.

The two projects are outlined in the modules shown in Fig. 8. Project 1 focuses on functional hand manipulation and spatial interaction and Project 2 emphasizes gesture-driven robotic control and kinematic reasoning. Each project was designed using the Magic Leap 2 AR platform and a general-purpose workstation running Python code to implement MediaPipe for hand movement analysis.

During the project experiments, continuous, timestamped hand joint data are collected in both AR and MediaPipe systems; IP, DIP, PIP, and MCP bone joint angles are computed from 3D hand bone data in AR and from relative 2D/3D landmark data in MediaPipe. These measurements enable the classification

Module 7: Project 1 – 3×3 Grid Object Placement (AR) Project Description Gesture Mapping (Grab / Release) Grid Logic & Object Placement AR Implementation Details Implement grabbing and releasing using hand gestures in AR100 pts	Module 9: Project 2 – Robotic Arm Control (AR) Robotic Arm Kinematics (Overview) Gesture-to-Joint Mapping AR-Based Control Logic MQTT - communication protocol Control virtual robotic arm joints using hand gestures100 pts
Module 8: Project 1 – 3×3 Grid Object Placement (MediaPipe) Gesture Detection Logic Screen-to-Grid Mapping Object Placement Algorithm MediaPipe Implementation Implement 3×3 grid placement using MediaPipe100 pts	Module 10: Project 2 – Robotic Arm Control (MediaPipe) Gesture Recognition for Each Joint Mapping Hand Gestures to Joint Angles MediaPipe Control Pipeline Control robotic arm using MediaPipe hand tracking100 pts

Figure 8. Canvas Project Module Sequence.

of static and dynamic hand gestures and support quantitative analysis of hand movement. Both AR and CV applications expose learners to clinically motivated hand movements, gesture recognition logic, and system-level tradeoffs between immersive three-dimensional sensing and the more accessible two-dimensional vision-based tracking.

A. Project 1: Gesture-based object manipulation in a structured grid

This project introduces students to hand gesture recognition, object manipulation, and spatial mapping through a simple but clinically relevant task. The activity is inspired by functional hand rehabilitation exercises that require grasping, transporting, and releasing objects with positional accuracy. From an engineering perspective, the task emphasizes gesture classification, coordinate mapping, and interaction state management.

In this hand manipulation task, users place virtual spheres or cubes into a three-by-three grid using grasp and release gestures. A closed hand gesture was coded to indicate grasping, while an open hand gesture triggered the release of the virtual object. This task was designed to highlight differences in spatial perception between AR and vision-based interaction. In the AR environment, objects existed in a three-dimensional space, whereas in the vision-based system, interaction occurred through a two-dimensional projection. By completing the task on both platforms, participants were able to directly compare depth perception, interaction precision, and system responsiveness.

1). Clinical and rehabilitation objective

This task reflects common upper-extremity rehabilitation activities used in PT, such as reach, grasp, transport, and release exercises designed to improve fine motor control and hand-eye coordination. The structured grid emphasizes spatial accuracy and goal-directed movement, aligning with clinical assessments and therapeutic tasks that focus on precision and endpoint control. For health science learners, the activity provides insight into how functional hand movements are captured digitally and how interaction constraints and feedback mechanisms can support motor learning and rehabilitation-

oriented task design.

2). AR-based implementation

In the AR implementation shown in Fig. 9, the grid is anchored in the physical environment, allowing interaction within a three-dimensional physical space. The system performs real-time hand tracking and gesture recognition using the headset's sensor-fusion capabilities. When a grab gesture is detected, the virtual object is parented to the tracked hand and follows the motion of the user's hand in space. Upon release, the system computes the nearest grid cell in world coordinates and places the object accordingly. This implementation highlights depth perception, spatial accuracy, and world-locked interaction.

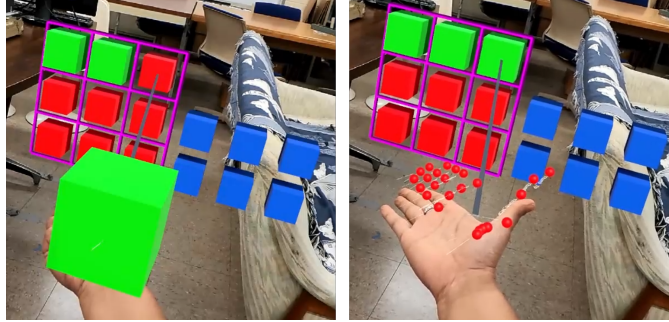


Figure 9. AR-based hand tracking interface showing object manipulation within a 3x3 grid. The left image shows the user grasping a blue cube at grid cell 3 with a closed hand; the corresponding blue cube turns green and attaches to the hand and targets the red grid cell 3. The right image shows the user releasing the cube with an open hand over the same cell, placing the blue cube onto the red cube. When the number of blue cubes correctly placed on red cubes matches the target, the red cubes turn green; otherwise, they remain red, providing visual feedback on task accuracy.

3). Mediapipe-based implementation

In the MediaPipe implementation shown in Fig. 10, hand landmarks are detected from a monocular webcam feed, and the interaction occurs in a two-dimensional screen space. The three-by-three grid is defined as a logical division of the screen, with each cell corresponding to a fixed region. A grab gesture is inferred from the relative distances between the fingertips and the center of the palm, whereas a release gesture corresponds to the extension of the fingers. The centroid of the hand is mapped to the screen coordinates to determine the active grid cell at the moment of release. This implementation emphasizes accessibility, computer-vision-based gesture inference, and the limitations of depth perception in 2D systems.

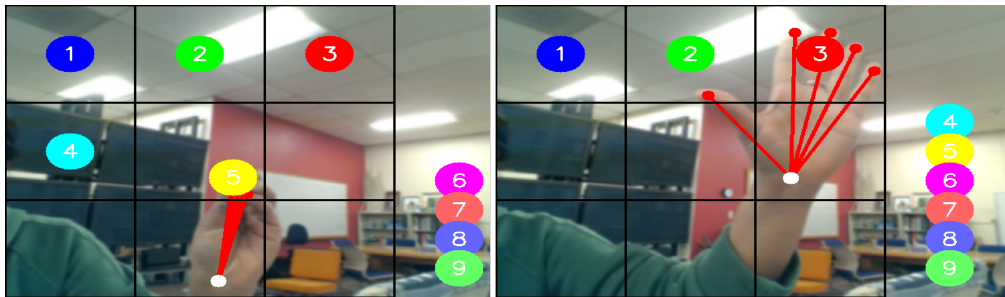


Figure 10. MediaPipe hand tracking interface showing a user performing grab and release gestures within a 3x3 spatial grid. The left image depicts the hand in a grabbing (close hand) pose targeting grid cell 5, while the right image shows the releasing (open hand) pose targeting grid cell 3. Colored circles correspond to numbered grid cells for interaction feedback.

B. Project 2: Gesture controlled robotic arm kinematics

The second project introduces learners to robotic kinematics, joint isolation, and human-robot interaction through gesture-based control of a multi-degree-of-freedom robotic arm. The activity emphasizes inverse kinematics concepts, joint constraints, and the importance of unambiguous gesture-to-command mapping. Clinically, the task reinforces controlled joint motion within defined ranges, a core principle in therapeutic exercise design.

This project addresses the control of a five-degree-of-freedom robotic arm using distinct hand gestures, each gesture mapped to a specific robot-joint, which includes base rotation, shoulder, elbow, wrist rotation, and gripper actuation. In both environments, a visual arc representing the range of the joint was displayed, with a sphere moving along the arc to indicate the joint angle based on hand-gesture movement. Joint motion is constrained within safe minimum and maximum limits, typically between 0 and 180 degrees, reflecting both mechanical limitations and clinically relevant ranges of motion. This task reinforced concepts of robotic kinematics, joint isolation, and human-robot interaction for engineering students, while emphasizing joint control and movement limits relevant to PT education.

1). Clinical and rehabilitation objectives

This project aligns with the rehabilitation principles of isolated joint control and constrained range-of-motion training commonly used in therapeutic exercise programs. Mapping distinct hand gestures to individual robotic joints reinforces selective motor activation and mirrors clinical strategies aimed at reducing compensatory movement patterns. For health science students, the task illustrates how human joint kinematics can be translated into controlled mechanical motion, supporting an understanding of rehabilitative robotics, biofeedback, and human-machine interfaces in clinical and assistive technology contexts.

2). AR-based implementation

In the AR environment shown in Fig. 11, the robotic arm model is world-locked and visualized in three dimensions. Hand gestures are detected using three-dimensional hand tracking, enabling intuitive control of joint motion through spatial hand movement. A visual arc representing the allowed range of motion of the joint is displayed, with a sphere moving along the arc to indicate the current angle of the joint. The detected joint angles are transmitted in real time to a physical or simulated robotic system using the Message Queuing Telemetry Transport (MQTT) communication protocol, which supports low-latency control and reinforces the concepts of system-integration.

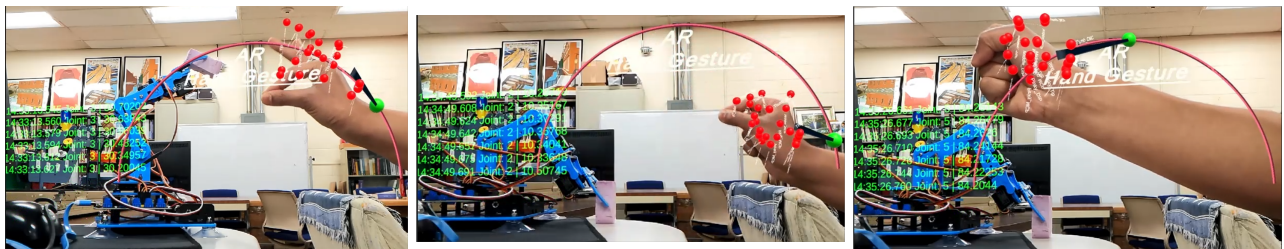


Figure 11. Augmented reality hand gesture control of a robotic arm. The hand center is connected by a gray line to a green-sphere on the arc, which represents joint angle values ranging from 0° to 180°. The first image depicts joint 3 positioned at approximately 30° controlled by the “OK” gesture, the second image shows joint 2 at about 10° with a pinch gesture, and the third image illustrates joint 5 at roughly 84° using a closed hand gesture. In all three images, the robotic arm holds and moves a cube, demonstrating how distinct hand gestures correspond to specific joint movements for manipulating objects. The green-text represents the real-time robotic-arm joint number and it's angles with timestamp.

3). Mediapipe-based implementation

In the MediaPipe-based system depicted in Fig. 12, gesture recognition relies on the relative positions and angles of the 21 detected hand landmarks. Specific static gestures are used to select individual joints, while horizontal or vertical hand movement adjusts the joint angle. The same arc-and-sphere visualization is rendered in screen space to provide consistent feedback across platforms. This implementation emphasizes gesture classification, motion analysis, and the challenges of controlling continuous joint variables using two-dimensional hand tracking.

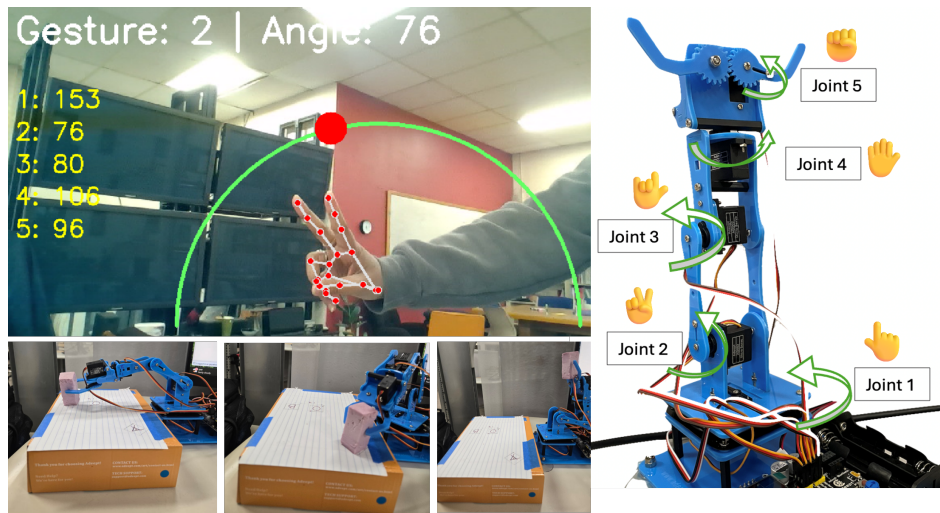


Figure 12. Hand gesture-controlled robotic arm system using web-camera input and MediaPipe for hand tracking, enabling live interaction and joint control, demonstrated by a user performing tasks such as moving objects from one target-place to another.

C. Educational goals

Together, these two projects provide complementary learning experiences that integrate hand tracking, gesture recognition, and system-level reasoning. By implementing identical tasks across AR and vision-based platforms, learners gain practical insight into how sensing modality, dimensionality, and hardware constraints influence system design decisions.

During experiment execution, the systems continuously track the joint spatial positions and angles of all ten fingers with timestamps. The interaction and task-related data are also continuously logged and stored. These data logs enable post hoc analysis, cross-platform comparison, and validation against clinical reference measurements.

V. Discussion

The AR-CV research and educational frameworks were developed and evaluated in close collaboration with faculty advisors, graduate students from engineering and PT. The PT researchers contributed to defining joint-level metrics, refining hand-gesture classification thresholds, and interpreting range-of-motion data generated by both AR and MediaPipe systems. Preliminary feedback indicates that automated, timestamped joint-angle logging reduces manual documentation effort while improving speed, consistency, and reliability compared to traditional goniometer-based assessments. The collected datasets are currently supporting ongoing clinical research and curriculum development in PT education.

By comparing immersive AR sensor-fusion systems with accessible computer-vision-based solutions, students engage with system-level tradeoffs in accuracy, latency, depth perception, scalability, and cost. The framework promotes active and reflective learning through gesture-driven system design, kinematic data logging, and cross-disciplinary collaboration grounded in clinically informed metrics.

References

- [1] G. Amprimo, G. Masi, G. Pettiti, G. Olmo, L. Priano, and C. Ferraris, "Hand tracking for clinical applications: Validation of the google mediapipe hand (gmh) and the depth-enhanced gmh-d frameworks," *Biomedical Signal Processing and Control*, vol. 96, p. 106508, 2024.
- [2] R. Husna, K. C. Brata, I. T. Anggraini, N. Funabiki, A. A. Rahmadani, and C.-P. Fan, "An investigation of hand gestures for controlling video games in a rehabilitation exergame system," *Computers*, vol. 14, no. 1, 2025.
- [3] A. I. Stoumpos, F. Kitsios, and M. A. Talias, "Digital transformation in healthcare: Technology acceptance and its applications," *International Journal of Environmental Research and Public Health*, vol. 20, no. 4, 2023.
- [4] M. Mauro, G. Noto, A. Prenestini, and F. Sarto, "Digital transformation in healthcare: Assessing the role of digital technologies for managerial support processes," *Technological Forecasting and Social Change*, vol. 209, p. 123781, 2024.
- [5] B. J. Jo, S.-K. Kim, and S. Kim, "Enhancing virtual and augmented reality interactions with a mediapipe-based hand gesture recognition user interface," *Ingénierie des Systèmes d'Information*, vol. 28, no. 3, pp. 633–638, 2023.
- [6] J. Grodotzki, B. T. Müller, and A. E. Tekkaya, "Enhancing manufacturing education based on controller-free augmented reality learning," *Manufacturing Letters*, vol. 35, pp. 1246–1254, 2023, 51st SME North American Manufacturing Research Conference (NAMRC 51).
- [7] F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G. Sung, C. Chang, and M. Grundmann, "Mediapipe hands: On-device real-time hand tracking," *CoRR*, vol. abs/2006.10214, 2020.
- [8] S. G. Létourneau, H. Jin, E. Peters, R. Grewal, D. Ross, and C. Symonette, "Augmented reality-based finger joint range of motion measurement: Assessment of reliability and concurrent validity," *The Journal of Hand Surgery*, vol. 50, no. 8, pp. 966–978, 2025.
- [9] G. Krause-Jüttler, J. Weitz, and U. Bork, "Interdisciplinary collaborations in digital health research: Mixed methods case study," *JMIR Human Factors*, vol. 9, no. 2, p. e36579, May 2022.
- [10] B. N. Green and C. D. Johnson, "Interprofessional collaboration in research, education, and clinical practice: Working together for a better future," *Journal of Chiropractic Education*, vol. 29, no. 1, pp. 1–10, 2015.
- [11] W. Li and N. Su, "Innovative research on interdisciplinary collaborative education mechanisms driven by artificial intelligence," in *2025 5th International Conference on Artificial Intelligence and Education (ICAIE)*, 2025, pp. 666–669.
- [12] C. Budzinski, H.-L. Wu, E. Sarraf, S. Miller, and J. Moore, "Accuracy of mediapipe visual hand tracking for use in medical training procedures," *Proceedings of the ASME 2024 Design of Medical Devices Conference*, vol. V001T08A004, Apr. 2024.
- [13] J.-S. Han, C. I. Lee, Y. H. Youn, and S. J. Kim, "A study on real-time hand gesture recognition technology by machine learning-based mediapipe," *Journal of System and Management Sciences*, vol. 12, no. 2, pp. 468–482, 2022.
- [14] M. Z. Iqbal and A. G. Campbell, "Real-time hand interaction and self-directed machine learning agents in immersive learning environments," *Computers Education: X Reality*, vol. 3, p. 100038, 2023.
- [15] T. J. Eiler, V. Schmücker, C. Gießer, and R. Brück, "Virtual and augmented reality for digital medicine - design and implementation of a hybrid, interdisciplinary course for engineering and medical students," *IEEE Global Engineering Education Conference (EDUCON) 2023*, 2023.

- [16] M. Connolly, G. Iohom, N. O'Brien, J. Volz, A. O'Muircheartaigh, P. Serchan, A. Biculescu, K. G. Gadre, C. Soare, L. Griseto, and G. Shorten, "Delivering clinical tutorials to medical students using the microsoft hololens 2: A mixed-methods evaluation," *BMC Medical Education*, vol. 24, no. 1, p. 498, 2024.
- [17] O. Zechner, H. G. Schrom-Feiertag, R. Wespi, D. Pretolesi, H. Q. H. Nguyen, and M. Tscheligi, "Enhancing mixed reality simulation training technology with real-time performance visualization: Mixed methods study with medical first responders," *JMIR XR and Spatial Computing (JMXR)*, vol. 1, p. e57655, 2024.
- [18] B. Alsharif, E. Alalwany, A. Ibrahim, I. Mahgoub, and M. Ilyas, "Real-time american sign language interpretation using deep learning and keypoint tracking," *Sensors*, vol. 25, no. 7, 2025.
- [19] N. Suhail, Z. Bahroun, and V. Ahmed, "Augmented reality in engineering education: enhancing learning and application," *Frontiers in Virtual Reality*, vol. 5, p. 1461145, 2024.
- [20] Z. Lei, H. Zhou, W. Hu, and G.-P. Liu, "Web-based digital twin online laboratories: Methodologies and implementation," *Digital Twin*, vol. 2, p. 3, 2024.
- [21] G. Boopathy, E. Aoki, F. Norcéide, C. Thompson, E. Lewis, and K. Chandra, "Digital engineering framework for promoting health and human performance with immersive technology," *Proceedings of the IEEE International Systems Conference (SysCon)*, pp. 1–8, 2025.
- [22] G. Boopathy, "Hand tracking and gesture classification using augmented reality technology and machine learning algorithms," Master's thesis, University of Massachusetts Lowell, 2025, proQuest Dissertations & Theses Global, Document ID: 31934067.
- [23] J. Buckley and C. Exton, "Bloom's taxonomy: a framework for assessing programmers' knowledge of software systems," in *11th IEEE International Workshop on Program Comprehension, 2003.*, 2003, pp. 165–174.