Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 136 (2025) 377–382

www.elsevier.com/locate/procedia

35th CIRP Design 2025

Smart robopatients: Evolving the designs of virtual reality-based patient simulators for clinical training

Amar Kumar Behera[a,*]

[a]*Department of Design, Indian Institute of Technology, Kanpur 208016 India.*

* Corresponding author. Tel.: +91 512 259 2401 *E-mail address:* amarkb@iitk.ac.in

Abstract

A novel framework was developed to create a self-learning smart robopatient for clinical training that adapts its morphology and behavior with time, based on field data from clinical cases. A set of three functional modalities of the robopatient, viz.: patient assessment, monitoring and treatment, are available within a virtual reality (VR) environment. The entities within each modality are dynamically updated based on automated analysis of patient-doctor/nurse interaction data available as videos, prescriptions, notes, etc., creating a robust process flow for improving the design of the VR-based patient simulator with time. This, in turn, enables continually improved training of medical professionals.

Peer-review under responsibility of the scientific committee of the 35th CIRP Design 2025

Keywords: patient simulator; virtual reality; robopatient; clinical education; automated analysis

1. Introduction

The use of virtual patients in clinical training has gained significant popularity in recent years. These interactive, computer-based simulations or physical prototypes integrated with computer-generated imagery enable experiential learning and assessment in a safe, standardized environment. By simulating real-life clinical scenarios, virtual patients enable trainee doctors and healthcare practitioners to practice obtaining patient histories, conducting physical examinations, and making diagnostic and therapeutic decisions.

The term "virtual patient" encompasses various forms of simulation, including still images, video animations, avatars, and immersive virtual reality environments [1]. The fidelity of these representations has steadily improved through ongoing research and development in both academic and industry settings. As a result, 3D animations and virtual reality simulations have become increasingly realistic, leading to the emergence of "robopatients" [2]. These can be either asynchronous and accessible anytime and providing responses without a live tutor or synchronous providing real-time interaction with multiple users or tutors.

Simulation-based clinical education has gained further traction in recent years due to its higher effectiveness compared to traditional teaching methods involving standardized patient actors, physical mannequins, and cadavers [3]. One of the main advantages is to confront students with their mistakes enabling them to learn in the process. Commercially available robotic patients such as SimMan3G, Apollo CAE and Paediatric HAL are capable of responding to physical inputs and providing movement, verbal and haptic feedback to the student [4]. These robots can also display facial expressions like pain and discomfort, enhancing the realism of the training experience.

While certain mechanical robots such as Erica and Sophia can mimic human facial expressions using actuated motions, they are limited in their ability to represent faces from diverse demographics [5]. On the other hand, computer graphics-based simulation systems such as FACSHuman can render faces from different demographics [6], but they lack the ability to respond to physical inputs.

Peer-review under responsibility of the scientific committee of the 35th CIRP Design 2025
10.1016/j.procir.2025.08.066

2. Background and motivation

Current research focuses on high-fidelity representations that combine physical and virtual elements to create robopatients that can be programmed. A key limitation lies in the restricted range of shape representations currently feasible, which are mostly limited to facial transformations. Furthermore, only a small number of patient behaviours and medical conditions can be simulated and programmed in existing robopatient designs.

To address the limitations of physical and virtual representations of patients, hybrid (physical-virtual) systems have emerged. These typically use a fixed physical shell, with the virtual human projected onto the shell from the front or rear [7]. A major limitation of these hybrid systems is the mismatch caused by projection mapping for patients from different demographics. Since the rendered face is a key feedback modality, such mismatches reduce the fidelity of providing a realistic patient training experience. To resolve this issue, attempts have been made to morph the physical face to match the projected virtual face [3]. In another key development, a three-layered robotic mannequin shoulder has been created that can adjust the anisotropic deformation of its human-like skin to imitate body dimensions [8]. This design was developed by analysing the shape differences of numerous scanned human models. Despite this achievement, the researchers note that fully deformable and shape-adaptable robotic mannequins still require further work.

Additionally, current patient simulators still cannot fully display realistic conditions for training and assessment of competency including critical physical examination findings, such as breathing, mental condition, etc. and mimicry of clinical emergency events. It has also been reported that "visually-induced motion sickness," manifested as nausea, headaches and dizziness, may have a disruptive impact on some students at a physical level [9]. Recently, applications such as VRPatients [10], ORamaVR [11] and Simtoma [12] have been introduced to enhance training through gamification and VR integration. However, these platforms do not offer automated updates to clinical scenarios.

Apart from achieving perfect shape representations, efforts have also been made to realize patient-doctor interactions with increasing levels of fidelity. For instance, Bracegirdle et al. [13] developed a system in which a virtual patient utilizes a decision tree based on published clinical experience and evidence to respond to new questions, display 3D animations, and provide audio responses to the learner. Recent advancements have enabled AI models, such as ChatGPT, to pass the United States Medical Licensure Exam (USMLE) without specialized training [14]. However, when tested with real-world symptoms, its responses tend to be formulaic, heavily qualified, and similar to reading directly from WebMD. As such, AI tools remain far from being capable of replacing doctors.

As robopatients are highly individualized tools with significant design costs, their acceptability and usability have also come under scrutiny. The increased use of technology in professional healthcare education, especially following COVID-19 where students were compelled to engage in virtual education in various disciplines [15], has necessitated the drive to make robopatients more accessible to medical schools. The key to overcoming acceptance barriers for high-fidelity simulation devices in clinical education is threefold: (i) addressing scepticism, (ii) improving communication regarding their availability, and (iii) demonstrating their efficacy [16].

Given the current limitations of virtual patients in the medical sector, this research explores the development of a self-learning robopatient that adapts its morphology and behavior over time based on usage analysis of real-world data. A virtual environment was created with three functional modalities of the robopatient: patient assessment, monitoring, and treatment. A methodology was developed to update the entities within each modality through automated analysis of patient-doctor/nurse interaction data. This approach aims to establish a robust process for continuously improving the design of VR-based patient simulators.

3. Methodology

An integrated process flow was developed within this research that enables the seamless updating of a robopatient present within a virtual world. The robopatient is placed in a virtual world created using the commercial software package Unity, with relevant animations added to simulate specific clinical scenarios. The user interacts with a virtual interface that features three main menu items at the top: (i) Assess, (ii) Monitor, and (iii) Treat. Each menu item contains sub-menus that become visible upon selection. At the start of a session, the clinical trainee is presented with a specific scenario and must navigate the user interface to analyse the situation and make appropriate decisions. The trainee's choices are recorded in a database and sent to the clinical instructor for evaluation and feedback.

To update the virtual world, real-world data from day-to-day clinical evaluations is processed based on the type of data. Video data is processed using a speech processing routine within Google Colaboratory that employs Whisper AI to convert the speech within the video to a text file. Digital prescriptions are processed as is to text files using the in-built converter within the PDF readers such as Adobe Professional while written prescriptions are processed using optical character recognition (OCR) using Sider AI to generate a text transcript. Next, the text file(s) from the real-world data are classified into ''Assess'', ''Monitor'' and ''Treat'' entities and then compared to the menu options in the database of the virtual clinical scenario to come up with suggestions for improvement. If these suggestions are accepted, the virtual environment is updated within Unity to create an updated definition for the robopatient and the associated clinical scenario, typically in the form of new menu items or use of an updated animation for the robopatient. This entire process flow is illustrated in Fig. 1.

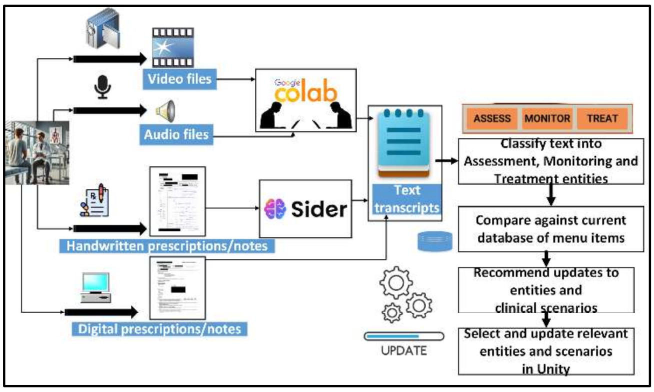


Fig. 1. Schematic for intelligent updating of clinical scenarios and entities based on real world clinical evaluations

A desk-based case study was developed to demonstrate the process flow for evolving the design of the virtual patient. This case study involves a hypothetical patient suffering from respiratory distress and abdominal pain. The patient is in his bedroom, and the call for assistance was placed by a neighbour who had not seen him for several days and had overheard bouts of coughing through the shared bedroom wall. The user is presented with an initial scene where the patient is in his bedroom, visibly coughing. Using a user menu placed next to the patient in the virtual environment, the user must decide on the appropriate next steps.

4. Implementation and results

The results of developing the virtual environment and implementing the process flow for smart robopatients are presented below.

4.1. Virtual world creation

A virtual patient application, ClinSimVR, was developed using the Universal Render Pipeline (URP) in Unity version 2022.3.8f1. A bedroom scene was created using assets from the Unity Asset Store, featuring a male character. A canvas with a top menu containing three tiers—Assess, Monitor, and Treat—was added to the side of the character, as shown in Fig. 2. A sitting and yelling animation from the Mixamo library [17] was imported and modified to simulate the appearance of an individual coughing.

Fig. 2. Opening scene in the virtual world presenting a patient and a canvas with three top menu items

Sub-menu items were displayed on user clicks, as shown in Fig. 3. Upon selecting the final sub-menu item in the hierarchy, the user received a status message about the patient's condition related to that item. This feature helps the user assess or monitor the patient's state, influencing subsequent decision-making within the scenario.

The key to developing a user-friendly canvas is to arrange menu items as clickable buttons in a neat and orderly manner while considering the canvas dimensions, text font size, and visibility of relevant items based on user interactions. This approach is also crucial for future updates based on real-life scenarios, as new menu items may need to be added or old ones removed while maintaining an organized and legible interface.

The XR Interaction Toolkit was used to integrate Meta Quest 2 controls into the project, allowing users to navigate the menu using VR controllers. The virtual world was tested within Unity and functioned as expected, enabling users to progress through the scenario and make informed choices to handle the clinical situation.

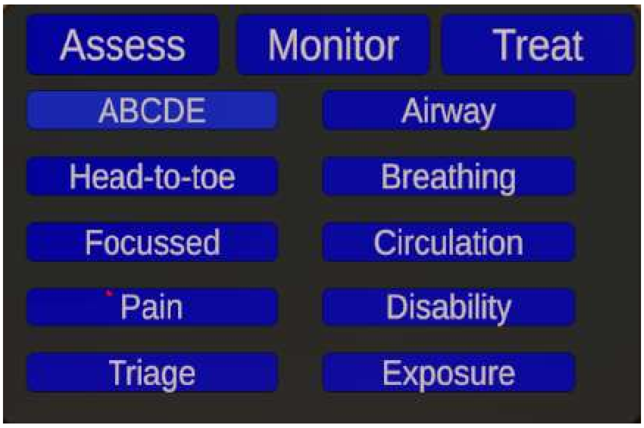


Fig. 3. Sub-menu items within the ClinSimVR app showing expansion of Assess and further expansion of ABCDE into five sub-items each

4.2. Video and audio transcribing

Google Colaboratory and Whisper AI were used to efficiently transcribe video files to text. First, Whisper AI was installed in the Colaboratory environment using commands to retrieve it from GitHub and install the necessary dependencies, including FFmpeg [18] for handling audio and video files. After installation, the transcription process was initiated by running Whisper with the command !whisper "filename" --model medium, which utilized the medium-sized model for optimal transcription accuracy. Additional arguments were available with the !whisper -h command, allowing customization of the transcription process, such as model size or language options, enhancing the tool's flexibility for various transcription needs. This setup enabled rapid, accurate text conversion from spoken content in videos, facilitating further analysis and processing.

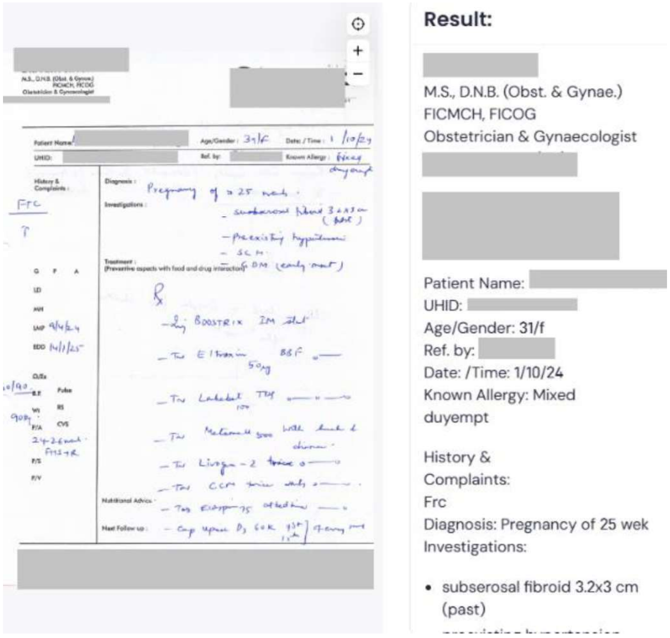
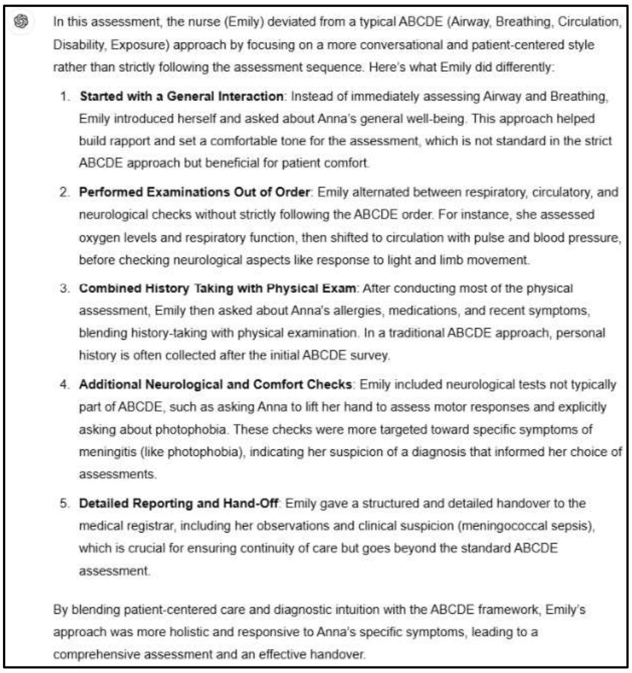


Fig. 4. Optical character recognition using Sider AI on a handwritten prescription

4.3. Prescription and notes parsing

Sider AI [19] was employed to convert doctor prescriptions and clinical notes from PDF format into editable text documents. It employs advanced Optical Character Recognition (OCR) technology to accurately interpret and transcribe both printed and handwritten text within PDF files, extracting essential medical information for digital use (see Fig. 4 for an example of a processed prescription). Upon uploading PDF prescriptions and notes, Sider AI processes each document, capturing the text data and making it accessible for analysis, storage, and integration into the virtual patient management system.



In this assessment, the nurse (Emily) deviated from a typical ABCDE (Airway, Breathing, Circulation, Disability, Exposure) approach by focusing on a more conversational and patient-centered style rather than strictly following the assessment sequence. Here's what Emily did differently:

1. **Started with a General Interaction**: Instead of immediately assessing Airway and Breathing, Emily introduced herself and asked about Anna's general well-being. This approach helped build rapport and set a comfortable tone for the assessment, which is not standard in the strict ABCDE approach but beneficial for patient comfort.
2. **Performed Examinations Out of Order**: Emily alternated between respiratory, circulatory, and neurological checks without strictly following the ABCDE order. For instance, she assessed oxygen levels and respiratory function, then shifted to circulation with pulse and blood pressure, before checking neurological aspects like response to light and limb movement.
3. **Combined History Taking with Physical Exam**: After conducting most of the physical assessment, Emily then asked about Anna's allergies, medications, and recent symptoms, blending history-taking with physical examination. In a traditional ABCDE approach, personal history is often collected after the initial ABCDE survey.
4. **Additional Neurological and Comfort Checks**: Emily included neurological tests not typically part of ABCDE, such as asking Anna to lift her hand to assess motor responses and explicitly asking about photophobia. These checks were more targeted toward specific symptoms of meningitis (like photophobia), indicating her suspicion of a diagnosis that informed her choice of assessments.
5. **Detailed Reporting and Hand-Off**: Emily gave a structured and detailed handover to the medical registrar, including her observations and clinical suspicion (meningococcal sepsis), which is crucial for ensuring continuity of care but goes beyond the standard ABCDE assessment.

By blending patient-centered care and diagnostic intuition with the ABCDE framework, Emily's approach was more holistic and responsive to Anna's specific symptoms, leading to a comprehensive assessment and an effective handover.

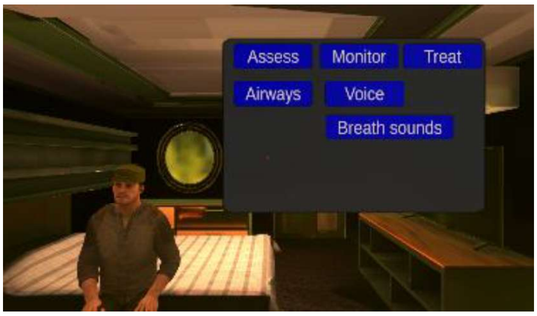
Fig. 5. Results of analysing the video text using ChatGPT

4.4. Updates within Unity

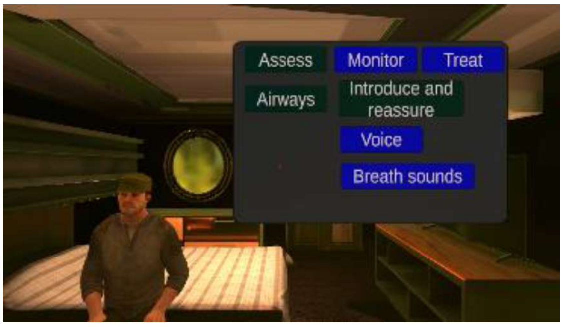
An ABCDE assessment video from the Dundee Emergency and Critical Care Society [20] was used to implement the workflow for updating the VR app within Unity. The video's text transcript was extracted using Whisper AI in Google Colaboratory and then analyzed with ChatGPT [21], an AI tool. ChatGPT was queried to identify deviations in the nurse's approach from the traditional ABCDE assessment framework. Next, it was asked to suggest updates to the menu items of the VR app (Fig. 5). For all ABCDE sub-menu items, the AI recommended retaining the existing options while adding a few additional ones based on the video analysis (Table 1). Following these suggestions, the menu and sub-menu items were updated, with the canvas resized and rearranged as needed (Fig. 6).

Table 1. AI recommendations for updating robopatient definition in Unity

ABCDE entity		Existing sub-menu item in VR app	AI recommendation
A-	Airways	Voice Breath sounds	Keep existing entities and add "Introduce and Reassure"
B-	Breathing	Respiratory rate Chest wall movements Chest percussion Lung auscultation Pulse oximetry	Keep existing entities
C-	Circulation	Skin color and sweating Capillary refill time Palpate pulse rate Heart auscultation Blood pressure Electrocardiography monitoring	Keep existing entities and add "Assess for shock signs"
D-	Disability	Level of consciousness – AVPU (alert, voice responsive, pain responsive or unresponsive) Limb movements Pupillary light reflexes Blood glucose	Keep existing entities
E-	Exposure	Expose skin Temperature	Keep existing entities and add "Check abdominal tenderness"
	--	--	Add additional menus for "Gather patient history" and "Symptoms and recent changes"



(a)

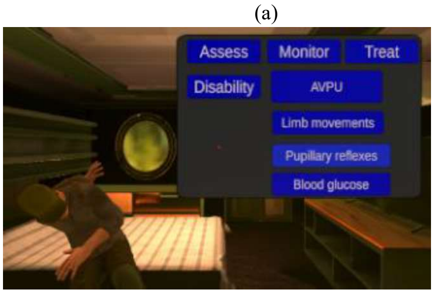


(b)

Fig. 6. Introducing new menu items in the ClinSimVR app showing (a) list in the original app and (b) list in the updated app

The AI tool was also queried for modifications to the robopatient's morphology and animations based on the nurse-patient interaction. One of the recommendations was to "Simulate the patient's reaction to light if they are photophobic, such as wincing or shielding their eyes, to enhance neurological assessment realism". This was implemented using two separate Mixamo animations - one depicting abdominal pain and coughing and the other illustrating wincing. When the pupillary reflexes button was clicked, a transition between the two animations was triggered using a trigger parameter in Unity (Fig. 7).

(a)



(b)

Fig. 7. Animation transitions implemented based on AI recommendations: (a) the top animation shows the patient coughing while (b) the bottom animation shows him shying away from light indicating photophobia

5. Discussion and conclusions

The self-learning smart robopatient framework presented in this study signifies a breakthrough in VR-based clinical training technology. By integrating adaptive, data-driven updates, this framework aims to create a highly realistic, responsive training environment that accurately mirrors real-world clinical complexity. The automated updating process is especially innovative, employing machine learning algorithms to analyze actual patient-doctor and patient-nurse interaction data, such as video interactions, prescriptions, and notes. This continual analysis refines the robopatient's behavior, morphology, and response options for the trainee student. This dynamic updating process enhances the simulation's fidelity and ensures the training remains relevant as clinical protocols evolve.

The framework's modular approach breaks down robopatient interactions into three categories: assessment, monitoring, and treatment. These categories mirror key stages of clinical practice, offering trainees a structured simulation experience. By interacting with these modules in a realistic setting, trainees gain an in-depth understanding of procedural and decision-making requirements. The adaptive design allows these modules to evolve based on new data, ensuring alignment with real-world practice standards and patient variations.

Using diverse data sources, including video analysis with Whisper AI, Sider AI OCR for handwritten notes, and digital prescriptions, strengthens the self-learning mechanism. This multi-source data processing ensures that the robopatient's responses are accurate and contextually relevant, incorporating nuances from various clinical encounters. Additionally, the data-driven approach fosters adaptability, seamlessly incorporating new trends or treatment protocols into the training environment. This ensures medical professionals are prepared for both standard scenarios and variations or unexpected patient responses.

While the framework shows significant strengths, it also faces challenges and areas for improvement. Integrating varied data types like video, audio, and text requires advanced machine learning and natural language processing (NLP) techniques to accurately interpret context and map it to VR interactions. Current NLP limitations, particularly with medical terminologies and nuances, may affect the precision of updates to the robopatient's behavior and morphology. Additionally, ethical and privacy concerns around using real patient data must be rigorously addressed to ensure data is de-identified and compliant with healthcare privacy standards.

A key limitation is automating the creation and management of new entities in game engines like Unity. This was done manually based on suggestions from ChatGPT. New entities, such as menu items within a canvas, required canvas resizing while considering the position of other game objects like the patient character. Introducing new entities also creates new logic for the training scenario, which needs to be programmed. This logic may not be readily apparent to an automation tool or AI, unlike human experience.

A key challenge also lies in updating the animations within the virtual world, which currently relies on a library of pre-defined animations. Creating custom animations based on actual video footage requires further exploration. Additionally, given the wide variety of human behaviors in specific situations—dependent on factors such as culture, age, gender, and personal experiences—there needs to be criteria for deciding which behaviors to include and which to exclude to provide a relevant training experience for trainee nurses and doctors.

6. Future work

In future developments, extending this framework to include more personalized patient responses based on demographic data or specific medical histories could enhance the realism of the simulation. Additionally, incorporating feedback loops with trainee performance data could provide insights into areas where medical students or professionals may need further training or guidance. Creating dynamic scenario scripting systems could allow instructors to adjust robopatient conditions, monitor changes, and assess treatment effectiveness in real-time.

It is also important to track and record students' actions and outcomes, logging correct and incorrect interventions. While this feature is available in current medical simulation apps, implementing it for new, evolving robopatient definitions requires expert judgement from educators, who may not always have clear patient outcomes or documented evidence for specific medical interventions. Expanding the framework to integrate wearable sensors or haptic feedback devices in VR could elevate the tactile experience of patient interactions, enriching hands-on learning.

Future work could involve conducting usability testing with clinical students and instructors to ensure scenarios are realistic, intuitive, and educational. Based on feedback, it will be useful to make necessary adjustments to improve the interface, flow of case logic, and scenario accuracy. It would also be worthwhile to create guides for instructors on editing scripts or database entries to create or modify new scenarios. Training instructors on using the VR app and the instructor interface will help maximize the effectiveness of smart rob patients.

In summary, this self-learning robopatient framework sets a foundation for continuously improving VR-based clinical training. By adapting to real-world clinical data and evolving patient-care practices, this approach holds promise for equipping medical professionals with the skills and confidence to handle complex, dynamic clinical scenarios.

Acknowledgements

This work was made possible by financial support from the Indian Institute of Technology, Kanpur [grant numbers R&D-DESP-2023284 and IITK-DESP-2023107].

References

[1] Thompson, J., White, S., and Chapman, S., 2020, "Interactive clinical avatar use in pharmacist preregistration training: Design and review," Journal of Medical Internet Research, 22(11), p. e17146.

[2] Tan, Y., Rérolle, S., Lalitharatne, T. D., Van Zalk, N., Jack, R. E., and Nanayakkara, T., 2022, "Simulating dynamic facial expressions of pain from visuo-haptic interactions with a robotic patient," Scientific Reports, 12(1), p. 4200.

[3] Lalitharatne, T. D., Tan, Y., He, L., Leong, F., Van Zalk, N., de Lusignan, S., Iida, F., and Nanayakkara, T., 2021, "Morphface: A hybrid morphable face for a robopatient," IEEE Robotics and Automation Letters, 6(2), pp. 643-650.

[4] Gudadappanavar, A. M., Hombal, P., Benni, J. M., Patel, S., and Tubaki, B. R., 2023, "Evaluation of Virtual Reality High-Fidelity Adult Mannequin-based Simulation of Real-Life Clinical Scenarios in Teaching Clinical Pharmacology to Medical Students," Journal of Pharmacology and Pharmacotherapeutics, p. 0976500X231159456.

[5] Lalitharatne, T. D., Costi, L., Hashem, R., Nisky, I., Jack, R. E., Nanayakkara, T., and Iida, F., 2022, "Face mediated human–robot interaction for remote medical examination," Scientific reports, 12(1), p. 12592.

[6] Gilbert, M., Demarchi, S., and Urdapilleta, I., 2021, "FACSHuman, a software program for creating experimental material by modeling 3D facial expressions," Behavior Research Methods, 53(5), pp. 2252-2272.

[7] Bermano, A., Brüschweiler, P., Grundhöfer, A., Iwai, D., Bickel, B., and Gross, M., 2013, "Augmenting physical avatars using projector-based illumination," ACM Transactions on Graphics (TOG), 32(6), pp. 1-10.

[8] Long, J., Li, J., Lu, Y., Zhou, C., Lu, G., and Feng, Y., 2023, "Shape-changing Robotic Mannequin Shoulder with Bio-inspired Layered Structure," IEEE Robotics and Automation Letters.

[9] Wu, Q., Wang, Y., Lu, L., Chen, Y., Long, H., and Wang, J., 2022, "Virtual simulation in undergraduate medical education: a scoping review of recent practice," Frontiers in Medicine, 9, p. 855403.

[10] VES, 2024, "VRPatients - Transform Your Educational Curriculum with Unmatched Realism," https://vrpatients.com/.

[11] ORamaVR, S., 2024, "Accelerating Computational Medical XR," https://oramavr.com/.

[12] Simtoma, 2024, "A powerful, practical, all-in-one clinical tool," https://simtoma.com/.

[13] Bracegirdle, L., and Chapman, S., 2010, "Programmable Patients: Simulation of Consultation Skills in a Virtual Environment," Bio Algorithms Med Syst., 6(11), pp. 111-115.

[14] DiGiorgio, A. M., and Ehrenfeld, J. M., 2023, "Artificial Intelligence in Medicine & ChatGPT: De-Tether the Physician," Journal of Medical Systems, 47(1), p. 32.

[15] Behera, A. K., de Sousa, R. A., Oleksik, V., Dong, J., and Fritzen, D., 2023, "Student perceptions of remote learning transitions in engineering disciplines during the COVID-19 pandemic: a cross-national study," European Journal of Engineering Education, 48(1), pp. 110-142.

[16] Rosen, K. R., 2008, "The history of medical simulation," Journal of critical care, 23(2), pp. 157-166.

[17] Mixamo, 2024, "Mixamo," https://www.mixamo.com/.

[18] FFmpeg, 2024, "A complete, cross-platform solution to record, convert and stream audio and video," https://www.ffmpeg.org/.

[19] Sider, 2024, "Sider: ChatGPT Sidebar with o1-preview & o1-mini now!," https://sider.ai/.

[20] Society, D. E. a. C. C., 2021, "DECCS ABCDE Video," https://www.youtube.com/watch?v=rVlmYb-2tyM.

[21] OpenAI, 2024, "ChatGPT," https://www.chatgpt.com.