ReMotion-AI: A Human-in-the-Loop Adaptive Control Framework for Compensation-Aware Rehabilitation Gaming

Shabnam Sadeghi Esfahlani
Science and Engineering
Anglia Ruskin University
Chelmsford, United Kingdom
ORCID: 0000-0002-1443-0330

Jason D. Taylor
Iterum Dynamics Ltd,
London, United Kingdom

***Abstract*—Rehabilitation systems often face barriers to adoption due to specialised hardware requirements and limited cross-platform delivery. We present ReMotion-AI, a gamified rehabilitation platform that has transitioned from a depth-sensor desktop application to a monocular, browser-deployed solution. The system uses a server-side inference pipeline that combines Human Mesh Recovery (HMR)-based full-body 3D pose and mesh recovery, SmoothNet for temporal stabilisation, and MediaPipe Hands for lightweight real-time hand-state estimation. A key engineering objective is backward compatibility and the inference output is mapped to a standardised joint format.**

To support safe and clinically interpretable adaptation, an Impairment and Compensation Adaptation Module (ICAM) is introduced based on interpretable rules and calibrated scoring. ICAM estimates movement capability, including range of motion (ROM) and smoothness and detects compensatory strategies such as trunk substitution, shoulder elevation and lower-limb compensations including hip hiking and altered gait patterns. The system is formulated as a human-in-the-loop adaptive control framework, in which ICAM acts as a supervisory controller with bounded adjustment.

Preliminary technical validation showed a median end-to-end latency of 180 ms, a 25–30% reduction in upper-limb trajectory jitter after temporal smoothing and stable supervisory adaptation with no abrupt difficulty changes observed under noisy monocular measurements. The proposed framework demonstrates that webcam-based pose estimation, temporal smoothing and rule-based compensation monitoring can be integrated into a browser-deployed platform.

***Index Terms*—Human-in-the-loop control, Markerless rehabilitation, Monocular 3D pose, Compensation detection, Adaptive feedback, Remote rehabilitation**

I. Introduction

Effective rehabilitation typically requires high-repetition practice with appropriate progression and timely feedback [1]–[3]. However, many patients experience barriers to sustained participation due to travel burden, limited access to supervised sessions and reduced motivation during long-term home programs [1], [4]–[6]. Serious games offer a promising route to increase engagement by embedding therapeutic movements into goal-directed tasks and feedback-driven play [7], [8]. However, many existing camera-based rehabilitation solutions rely on dedicated sensing hardware or platform-specific installations, which can limit accessibility, scalability and long-term maintainability [9]–[12]. A practical alternative is to deliver rehabilitation games through the web using a standard webcam, thereby improving accessibility and inclusivity by reducing hardware requirements and enabling participation across a wider population with internet-connected devices [13]–[15]. This shift, however, introduces two key technical and clinical challenges. First, monocular motion capture can be temporally unstable, leading to jitter that affects avatar control and the reliability of kinematic measures such as joint angles and range of motion (ROM) [16], [17]. Second, users may achieve in-game objectives through compensatory strategies such as trunk lean or shoulder elevation that bypass the intended movement, potentially reducing therapeutic value and increasing strain [18]–[20]. These considerations motivate an approach that is both technically robust and clinically interpretable, combining stabilised pose signals for interaction with a rule-based ICAM controller that discourages compensation while supporting safe progression [2], [16], [18]. Accordingly, this paper presents ReMotion-AI, a browser-based, hardware-agnostic rehabilitation platform enabling markerless interaction using a standard webcam. The system features a stabilised monocular 3D tracking pipeline for real-time kinematic estimation within a closed-loop interaction. It further incorporates an interpretable, compensation-aware supervisory controller with bounded, safety-constrained adaptation (ICAM). Finally, the system is formulated within a control-theoretic human-in-the-loop framework, with analysis of practical stability under measurement noise. The tracking stack integrates Human Mesh Recovery (HMR) 2.0-S for SMPL (Skinned Multi-Person Linear model)-based body reconstruction [21], [22], SmoothNet for temporal refinement [16] and MediaPipe Hands for robust hand-state estimation [23]. Unlike many existing rehabilitation games that focus primarily on engagement or sensing, this work explicitly integrates control-theoretic safety, interpretability and compensation awareness within a deployable and clinically aligned system.

Figure 1 shows a rehabilitation task in a ReMotion-AI

Fig. 1. ReMotion-AI gameplay scene with markerless full-body tracking and real-time feedback.

scene. The user performs reaching and lower-limb balance movements mirrored by the in-game avatar. The interface provides real-time feedback through joint-angle overlays, including shoulder and elbow kinematics, and task progress indicators. Target objects support goal-directed movement, while a webcam inset enables markerless pose capture for avatar control and performance logging during upper- and lower-limb exercises. The WebGL game displays the avatar and target objects together with joint-angle estimates and progression indicators.

II. RELATED WORK

Vision-based rehabilitation platforms increasingly combine motion tracking and gamification to support home-based therapy and improve patient engagement. Existing systems can broadly be divided into depth-sensor-based (RGBD) and monocular-camera-based (RGB) approaches. Early rehabilitation systems frequently relied on RGBD sensors such as Microsoft Kinect or Intel RealSense, which provide relatively accurate 3D motion capture. For example, Silva et al. developed a Unity3D Kinect V2 rehabilitation game with real-time feedback and progress tracking, reporting high user motivation in healthy participants [24]. Despite their effectiveness, such systems depend on dedicated hardware, limiting scalability, accessibility, and long-term deployment.

To address these limitations, recent research has shifted toward monocular RGB approaches that operate using standard webcams. Ang et al. evaluated a markerless motion-analysis pipeline based on OpenPose for estimating joint angles and angular velocities from home-recorded squat videos, demonstrating agreement with marker-based motion capture for that task [25]. Mohamed et al. proposed TOSHFA, a webcam-based telerehabilitation system for low back pain using MediaPipe pose estimation and gamified biofeedback, reporting encouraging engagement despite some usability challenges in pilot studies [26]. Similarly, Fontenele et al. extended the Rehabilite Game+ platform with hand-pose estimation for upper-limb rehabilitation, achieving strong usability and movement-recognition performance in evaluations with clinicians and healthy participants [27]. Collectively, these studies demonstrate the feasibility of webcam-based rehabilitation while highlighting persistent challenges related to motion stability, reliability, and clinically meaningful adaptation.

From a modelling perspective, monocular 3D human reconstruction methods combine 2D keypoint estimation with parametric body models [17], [21]. The SMPL model, for example, provides a compact representation of human body shape and pose that supports downstream kinematic analysis, including joint-angle and range-of-motion (ROM) estimation [22]. HMR-based architectures further extend this concept by directly regressing SMPL parameters from single RGB images while incorporating model-based priors, as demonstrated by SPIN (SMPL Optimisation IN the loop) [21]. However, frame-wise monocular reconstruction often suffers from temporal instability, especially under occlusion or uncommon poses. SmoothNet addresses this issue through temporal refinement of pose sequences, reducing noise while preserving natural motion dynamics and demonstrating strong transferability across datasets and estimators [16]. Systems such as VirTele demonstrate how home-based rehabilitation can be combined with remote monitoring and clinician-guided progression [28]. Nevertheless, many deployed solutions still depend on specialised hardware, reinforcing the need for lightweight, browser-accessible systems that operate with commodity cameras. A further challenge in technology-supported rehabilitation is the reduced level of clinical supervision, which can allow patients to complete tasks using compensatory movement strategies rather than therapeutically desirable motion patterns [19], [29], [30]. Existing rehabilitation games often focus primarily on task completion without explicitly modelling compensation-aware adaptation or interpretable progression mechanisms. ReMotion-AI addresses these gaps by combining a web-deployed monocular tracking pipeline with an interpretable Impairment and Compensation Adaptation Module (ICAM). The ICAM computes clinically motivated kinematic features from stabilised pose data and derives capability and compensation scores within a rule-based framework incorporating hysteresis and reliability gating. This enables conservative, transparent, and stable progression suitable for staged technical and clinical validation. The framework also builds upon earlier ReHabGame studies exploring machine-learning-based adaptation and outcome prediction [10], [11], extending these ideas into an auditable and compensation-aware control framework for full-body rehabilitation.

III. METHODOLOGY

A. Design

Four requirements guided the redesign of ReMotion-AI. First, we prioritised *accessibility and inclusivity* by removing reliance on specialised sensors and enabling use with a standard webcam. Second, we aimed for *cross-platform delivery* by deploying the game as a WebGL build. Third, we maintained *backward compatibility* by standardising pose outputs to the joint interface expected by the existing gameplay logic, thereby avoiding the need to reauthor avatar control and task mechanics. Finally, we ensured *clinical relevance* by incorporating explicit monitoring of ROM and compensatory movement patterns to support safer progression and clinician interpretability.

To meet these requirements, the platform follows a client–server architecture. The client is a Unity WebGL game running in the browser that captures webcam RGB frames and renders an avatar driven by inferred pose. Frames are streamed to a backend inference service, which returns compact pose packets containing 3D joints and joint rotations suitable for real-time control and logs session events for later analysis. Session summaries and progress metrics are then surfaced through a clinician-facing analytics layer, such as trend visualisations and exportable reports, to support remote review of performance over time. This separation enables incremental upgrades because the game consumes a standardised skeletal representation that is independent of the underlying pose model, allowing inference improvements to be introduced without changing core gameplay code.

Web deployment introduces practical constraints, including browser security limitations, restricted filesystem access and performance considerations for networking and rendering. We therefore adopt a lightweight WebSocket-based streaming approach that decouples real-time rendering on the client from model inference on the server. This design also simplifies controlled updates to the inference pipeline.

The server-side inference pipeline is executed on GPU-enabled hardware to ensure real-time performance. While HMR-based full-body reconstruction models are computationally demanding, the client–server architecture allows scalable deployment by offloading inference from the browser. This design ensures that computational complexity does not limit accessibility for end users while maintaining sufficient throughput for interactive rehabilitation tasks. To mitigate temporal instability in monocular estimates, SmoothNet is applied as a learned temporal refinement module that reduces jitter [16]. For hand interaction, MediaPipe Hands provides lightweight hand landmarks, from which we infer open or closed states using short hysteresis to prevent flicker [23]. The resulting body and hand outputs are mapped to a standardised joint list and coordinate conventions expected by the existing ReMotion-AI modules. This modular approach preserves legacy avatar control, scoring triggers and task logic while transitioning the sensing modality from depth to monocular RGB.

On top of the stabilised pose and hand streams, we introduce ICAM, a real-time decision layer with clinician-readable adaptation logic. ICAM transforms windowed kinematic signals into (i) capability estimates, such as ROM, smoothness and coordination, (ii) compensation indicators, such as trunk substitution and shoulder elevation and (iii) adaptation recommendations that parameterise the difficulty engine. The module is designed to be auditable and suitable for early-stage clinical calibration, enabling conservative progression policies that explicitly discourage compensatory strategies.

B. Human-in-the-Loop

We formulate ICAM as a human-in-the-loop feedback control system. The rehabilitation process is modelled as a closed-loop interaction between the user, the sensing stack and the adaptive controller. In this interpretation, the patient acts as the rehabilitation plant, the monocular tracking pipeline functions as an observer and ICAM acts as a supervisory controller that updates task difficulty based on interpretable motion features. We adopt a discrete-time formulation with sampling interval Δt and use t to denote discrete time steps throughout the control analysis. The control formulation is intended as a supervisory and local stability interpretation of the adaptation layer, rather than a full physiological model of human motor recovery.

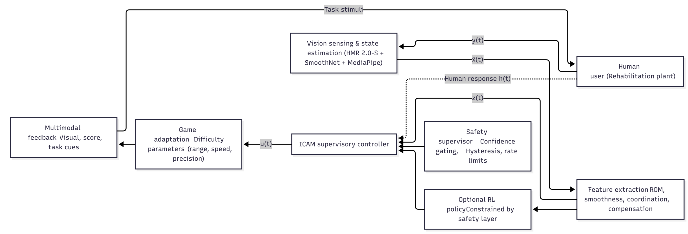

Fig. 2. Human-in-the-loop control architecture of ReMotion-AI.

Figure 2 shows the control-oriented interpretation of ReMotion-AI. Filtered full-body pose and hand-state streams are normalised and windowed to compute ROM, movement quality, coordination, compensation and hand-function features. These are combined into Capability (CS), Compensation (PS) and Control Confidence (CC) scores, which are passed through reliability gating, hysteresis and rate limiting to generate a stable reconfiguration signal for the difficulty engine and session-level analytics.

This formulation supports the application of established control-theoretic tools to analyse stability, boundedness and safety of the adaptation process, which are typically not explicitly addressed in rehabilitation gaming systems.

Let the latent rehabilitation state be

$$\mathbf{x}_t \in \mathbb{R}^n, \tag{1}$$

representing movement capability, compensation tendency and control reliability. The discrete-time closed-loop dynamics are written as

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \tag{2}$$

where

$$\mathbf{u}_t = [d_{\text{range}}, d_{\text{speed}}, d_{\text{precision}}]^\top \in \mathcal{U} \tag{3}$$

denotes the game adaptation input and $\mathbf{w}_t$ captures unmodelled human variability, fatigue and sensing disturbance.

The measured output is

$$\mathbf{y}_t = h(\mathbf{x}_t) + \mathbf{v}_t, \tag{4}$$

where $\mathbf{y}_t$ consists of pose-derived kinematic quantities and $\mathbf{v}_t$ is measurement noise.

From $\mathbf{y}_t$, a feature vector $\mathbf{z}_t$ is computed:

$$\mathbf{z}_t = [\text{ROM}_t,\ S_{\text{jerk},t},\ \rho_{\text{comp},t},\ \phi_{\text{trunk},t},\ C_{\text{stab},t}]^\top. \tag{5}$$

ICAM then computes the bounded adaptation command

$$\mathbf{u}_{t+1} = \Pi_{\mathcal{U}}\Big(\mathbf{u}_t + K\,\psi(\mathbf{z}_t)\Big), \tag{6}$$

where $\psi(\cdot)$ maps interpretable features to a signed progression or regression signal, K is a gain matrix and $\Pi_{\mathcal{U}}(\cdot)$ denotes projection onto the admissible control set defined by task safety limits.

C. Kinematic Features

The root-relative joint position is defined as

$$\tilde{\mathbf{p}}_j(t) = \mathbf{p}_j(t) - \mathbf{p}_r(t), \tag{7}$$

where $\mathbf{p}_j(t) \in \mathbb{R}^3$ is the 3D position of joint j at time t and $\mathbf{p}_r(t)$ is the root joint, such as the pelvis.

$$s(t) = \frac{1}{|\mathcal{B}|} \sum_{(a,b)\in\mathcal{B}} \|\tilde{\mathbf{p}}_a(t) - \tilde{\mathbf{p}}_b(t)\|_2, \qquad \hat{\mathbf{p}}_j(t) = \frac{\tilde{\mathbf{p}}_j(t)}{s(t) + \epsilon}, \tag{8}$$

where $\mathcal{B}$ is a set of limb segments and $\epsilon > 0$ avoids division by zero.

$$\theta(t) = \arccos\left(\frac{(\hat{\mathbf{p}}_a(t) - \hat{\mathbf{p}}_b(t)) \cdot (\hat{\mathbf{p}}_c(t) - \hat{\mathbf{p}}_b(t))}{\|\hat{\mathbf{p}}_a(t) - \hat{\mathbf{p}}_b(t)\|_2 \, \|\hat{\mathbf{p}}_c(t) - \hat{\mathbf{p}}_b(t)\|_2 + \epsilon} \right), \tag{9}$$

where b is the joint of interest and (a, b, c) are the proximal, joint and distal points.

$$\mathrm{ROM}_\theta(W) = \max_{t\in W} \theta(t) - \min_{t\in W} \theta(t). \tag{10}$$

$$\mathbf{v}(t) = \frac{\hat{\mathbf{x}}(t) - \hat{\mathbf{x}}(t - \Delta t)}{\Delta t}, \qquad \mathbf{a}(t) = \frac{\mathbf{v}(t) - \mathbf{v}(t - \Delta t)}{\Delta t}, \tag{11}$$

where $\hat{\mathbf{x}}(t)$ denotes the position of an end effector, such as the wrist or ankle.

$$\mathbf{j}(t) = \frac{\mathbf{a}(t) - \mathbf{a}(t - \Delta t)}{\Delta t}, \qquad S_{\mathrm{jerk}}(W) = \frac{1}{|W|} \sum_{t\in W} \|\mathbf{j}(t)\|_2^2. \tag{12}$$

$$\mathbf{q}_{\mathrm{trunk}}(t) = \frac{\hat{\mathbf{p}}_{\mathrm{thor}}(t) - \hat{\mathbf{p}}_r(t)}{\|\hat{\mathbf{p}}_{\mathrm{thor}}(t) - \hat{\mathbf{p}}_r(t)\|_2 + \epsilon}, \qquad \phi_{\mathrm{trunk}}(t) = \arccos(\mathbf{q}_{\mathrm{trunk}}(t) \cdot \mathbf{e}_z), \tag{13}$$

$\mathbf{e}_z = [0, 0, 1]^\top$ approximates the vertical axis in the normalised coordinate frame.

$$\rho_{\mathrm{comp}}(W) = \frac{\sum_{t\in W} \|\hat{\mathbf{p}}_r(t) - \hat{\mathbf{p}}_r(t - \Delta t)\|_2}{\sum_{t\in W} \|\hat{\mathbf{x}}(t) - \hat{\mathbf{x}}(t - \Delta t)\|_2 + \epsilon}. \tag{14}$$

$$C_{\mathrm{stab}}(W) = \exp\left(-\frac{1}{|W|} \sum_{t\in W} \sum_{j\in\mathcal{J}} w_j \left\|\hat{\mathbf{p}}_j(t) - \hat{\mathbf{p}}_j(t - \Delta t)\right\|_2^2 \right), \tag{15}$$

$\mathcal{J}$ is a joint set and w_j are joint weights, for example to emphasise the torso and target limb.

$$\begin{aligned} \mathrm{CS}(W) &= \sum_{k\in\mathcal{F}_{cap}} \alpha_k \bar{f}_k(W), \\ \mathrm{PS}(W) &= \sum_{k\in\mathcal{F}_{comp}} \beta_k \bar{f}_k(W), \\ \mathrm{CC}(W) &= \gamma \, C_{\mathrm{stab}}(W). \end{aligned} \tag{16}$$

where $\bar{f}_k(W)$ are normalised features over window W and $\alpha_k, \beta_k, \gamma$ are task-dependent weights.

$$g(W) \leftarrow \begin{cases} 1, & \mathrm{CC}(W) \geq \tau_{\mathrm{on}}, \\ 0, & \mathrm{CC}(W) \leq \tau_{\mathrm{off}}, \\ g(W^-), & \text{otherwise}, \end{cases} \qquad \text{with } \tau_{\mathrm{on}} > \tau_{\mathrm{off}}, \tag{17}$$

where W^- denotes the previous window.

$$A(W) = \begin{cases} \text{progress}, & g(W) = 1 \wedge \mathrm{CS}(W) \geq \tau_{cap} \wedge \mathrm{PS}(W) \leq \tau_{comp}, \\ \text{regress}, & g(W) = 1 \wedge \mathrm{PS}(W) > \tau_{comp}, \\ \text{hold}, & \text{otherwise}. \end{cases} \tag{18}$$

where τ_{cap} and τ_{comp} are calibrated per task family and applied with rate limiting in the controller. The decision variable in Eq. 18 determines whether the controller applies progression, regression, or hold actions through the bounded update law in Eqs. 19–21.

D. Feedback Mechanism

The ICAM controller operates as a supervisory feedback controller with safety constraints and hysteresis to prevent oscillatory behaviour. The bounded adaptation dynamics introduced in Eq. 6 can be expressed as

$$\mathbf{u}(t + 1) = \mathbf{u}(t) + \Delta\mathbf{u}(t), \tag{19}$$

where

$$\Delta\mathbf{u}(t) = K_p\big(\mathrm{CS}(t) - \tau_{cap}\big) - K_c\big(\mathrm{PS}(t) - \tau_{comp}\big), \tag{20}$$

subject to

$$\mathbf{u}_{\min} \leq \mathbf{u}(t) \leq \mathbf{u}_{\max}. \tag{21}$$

Hysteresis in Eq. 17 prevents rapid switching, while rate limiting constrains adaptation speed:

$$\|\Delta\mathbf{u}(t)\| \leq \delta_{\max}. \tag{22}$$

This design yields bounded adaptation updates under noisy pose estimates and supports practically stable closed-loop behaviour.

E. Adaptive Update

A desirable property of the adaptation mechanism is that game difficulty should not oscillate excessively in response to noisy measurements or transient compensatory motions. To analyse this behaviour, we consider the controller update error relative to a task-dependent equilibrium input $\mathbf{u}^\star$:

$$\mathbf{e}_t = \mathbf{u}_t - \mathbf{u}^\star. \tag{23}$$

The following analysis is intended to provide a local supervisory-control interpretation of the ICAM update rather than a formal experimental proof of human motor-system stability. The assumptions of local Lipschitz continuity and bounded disturbances are standard sufficient conditions used to reason about bounded adaptive updates, but they were not directly identified from participant-specific system models in this preliminary technical validation. Therefore, the analysis should be interpreted as a design-level argument showing how projection, hysteresis and rate limiting constrain the adaptation signal under bounded measurement perturbations.

This approximation reflects the assumption that increases in difficulty beyond the user's capability reduce performance scores, yielding a negative feedback relationship between the control error and feature response. Assume that, in a neighbourhood of the desired operating point, the feature-to-control map $\psi(\mathbf{z}_t)$ is locally Lipschitz and can be approximated as

$$\psi(\mathbf{z}_t) \approx -L\mathbf{e}_t + \mathbf{d}_t, \tag{24}$$

where L is positive semidefinite and $\mathbf{d}_t$ is a bounded disturbance term induced by sensing noise, temporary occlusions, and human variability. Substituting Eq. 24 into Eq. 6 gives the local closed-loop error dynamics

$$\mathbf{e}_{t+1} \approx \mathbf{e}_t - KL\mathbf{e}_t + K\mathbf{d}_t. \tag{25}$$

Consider the candidate Lyapunov function

$$V_t = \mathbf{e}_t^\top P \mathbf{e}_t, \qquad P = P^\top > 0. \tag{26}$$

For sufficiently small controller gains, such that the matrix $(I - KL)$ is Schur stable, the nominal system satisfies

$$V_{t+1} - V_t \leq -\mathbf{e}_t^\top Q \mathbf{e}_t, \qquad Q = Q^\top > 0, \tag{27}$$

in the absence of disturbances. With bounded disturbance $\|\mathbf{d}_t\| \leq \bar{d}$, the Lyapunov difference satisfies

$$V_{t+1} - V_t \leq -\lambda_{\min}(Q)\|\mathbf{e}_t\|^2 + c_1\|\mathbf{e}_t\|\bar{d} + c_2\bar{d}^2, \tag{28}$$

for some constants $c_1, c_2 > 0$. Under these assumed local conditions, the adaptation error is uniformly ultimately bounded.

IV. RESULTS

A technical validation study was conducted to evaluate real-time performance, signal stability and controller behaviour. A series of controlled end-to-end sessions was performed under representative operating conditions. A total of twelve sessions were conducted using internal engineering trials with fifteen healthy adult users performing predefined upper-limb and full-body movement tasks, including reaching, holding and controlled repetitions.

The client system consisted of a standard consumer laptop (Intel i7 CPU, integrated GPU, 16 GB RAM) running the Unity WebGL application in a modern web browser, while the server-side inference pipeline was deployed on a GPU-enabled backend (NVIDIA RTX-class GPU). Tests were conducted over a typical home broadband connection (Wi-Fi, 30–60 Mbps), without specialised network optimisation.

Performance metrics were computed from timestamped runtime logs, including frame-level inference time, client–server latency and rendering rate. Kinematic stability metrics, such as joint trajectory jitter and joint-angle variance, were derived from pose time-series data before and after temporal smoothing. Adaptation metrics, including false ROM-trigger events and difficulty oscillations, were quantified from threshold crossings and controller state transitions during gameplay segments.

Across repeated test sessions, the system sustained responsive interactive performance in the browser. The median end-to-end latency, from webcam frame acquisition to server-side inference, pose-packet transmission and client-side avatar update, was 180 ms (IQR: 150–240 ms), with a median rendering rate of 22 FPS in the Unity WebGL client. These results indicate that the client–server streaming design supports responsive gameplay under typical home-network conditions. Server-side inference time was measured on a GPU-enabled backend using per-frame processing logs, while round-trip latency was estimated from timestamped client–server message exchanges. False ROM-trigger events were quantified by counting spurious threshold crossings during controlled movement trials before and after temporal smoothing. These latency and rendering characteristics indicate responsive interaction for the tested deployment conditions.

We quantified temporal jitter using frame-to-frame variation in joint trajectories and joint-angle stability during static and controlled movements. Applying SmoothNet reduced the mean frame-to-frame displacement of key upper-limb joints by 25–30% and reduced elbow-angle variance during a static hold by 45% compared with the unsmoothed pose stream. This stabilisation improved avatar steadiness and reduced spurious threshold crossings in ROM-triggered game events.

We evaluated ICAM behaviour by analysing adaptation decisions over complete gameplay segments. With hysteresis and rate limiting enabled, difficulty oscillations were rare, with a median of 0–1 oscillations per session and progression was automatically suppressed when control confidence was low, with reliability gating active for 8% of windows. When confidence recovered and capability remained stable, ICAM produced gradual progression decisions consistent with the conservative safety policy. Controller credibility was assessed empirically through observable bounded-update behaviour, including difficulty oscillations, rate-limited progression/regression events, reliability-gating activity and the absence of abrupt difficulty changes across the tested sessions.

End-to-end deployment was validated across four browser and operating-system configurations. Session events and summary metrics were successfully stored and retrieved for dashboard review. From a control perspective, the bounded-update mechanism maintained stable adaptation trajectories across twelve test sessions, with low oscillation frequency and no abrupt difficulty jumps. This supports bounded supervisory adaptation under noisy webcam-based measurements.

V. DISCUSSION

The results indicate that ReMotion-AI can support responsive browser-based interaction and stable rule-based adaptation under controlled technical validation conditions. However, the present study should be interpreted as a preliminary technical evaluation rather than a clinical efficacy study. ICAM currently uses manually specified thresholds selected for conservative adaptation, and systematic parameter tuning, sensitivity analysis, and learning-based calibration were not performed.

The kinematic features used by ICAM, including ROM, smoothness, trunk angle, and compensation ratio, are established biomechanical descriptors. The contribution of this work is therefore their integration into a web-deployed, compensation-aware, human-in-the-loop adaptive rehabilitation framework rather than the introduction of new biomechanical metrics.

Several limitations remain. The validation was limited to healthy adult users and did not include patient cohorts or clinical outcome measures. Robustness to occlusion, lighting variation, camera viewpoint, and uncontrolled home environments was not systematically evaluated. Future work will compare ReMotion-AI-derived kinematics against reference motion-capture systems, tune ICAM thresholds using patient-specific data, and evaluate clinical outcomes in rehabilitation cohorts.

VI. Conclusions

This paper presented ReMotion-AI, a monocular WebGL-deployed rehabilitation game that integrates a server-side full-body inference pipeline based on an HMR-derived model, SmoothNet, and MediaPipe Hands, together with an interpretable Impairment and Compensation Adaptation Module (ICAM). The proposed architecture improves accessibility and scalability while supporting incremental system upgrades without disrupting existing gameplay logic through the use of standardised joint interfaces. By reframing ICAM as a human-in-the-loop adaptive control framework, the system enables explicit reasoning about feedback generation, bounded supervisory adaptation, and safety-aware constraints during rehabilitation tasks. In this way, the framework provides an initial step toward interpretable, compensation-aware, and bounded adaptive rehabilitation within a deployable real-time system.

The experimental results demonstrate the technical feasibility of the proposed framework under the evaluated conditions and indicate that the system is capable of satisfying practical real-time interaction requirements. Nevertheless, the present study does not establish clinical effectiveness, patient-level robustness, or superiority over alternative adaptive rehabilitation approaches, which remain important directions for future investigation.

References

[1] T. Kitago and J. W. Krakauer, "Motor learning principles for neurorehabilitation," *Handbook of clinical neurology*, vol. 110, pp. 93–103, 2013.

[2] J. A. Kleim and T. A. Jones, "Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage," *Journal of speech, language, and hearing research*, vol. 51, no. 1, pp. S225–S239, 2008.

[3] C. J. Winstein, J. Stein, R. Arena, B. Bates, L. R. Cherney, S. C. Cramer, F. Deruyter, J. J. Eng, B. Fisher, R. L. Harvey *et al.*, "Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the american heart association/american stroke association," *Stroke*, vol. 47, no. 6, pp. e98–e169, 2016.

[4] M. Kafri and O. Atun-Einy, "From motor learning theory to practice: a scoping review of conceptual frameworks for applying knowledge in motor learning to physical therapist practice," *Physical Therapy*, vol. 99, no. 12, pp. 1628–1643, 2019.

[5] National Institute for Health and Care Excellence (NICE), "Stroke rehabilitation in adults: Recommendations (NICE Guideline NG236)," https://www.nice.org.uk/guidance/ng236/chapter/Recommendations, 2023, accessed: 2026-03-04.

[6] A. Mahmood, P. Nayak, A. Deshmukh, C. English, J. Solomon *et al.*, "Measurement, determinants, barriers, and interventions for exercise adherence: A scoping review," *Journal of Bodywork and Movement Therapies*, vol. 33, pp. 95–105, 2023.

[7] C. S. González-González, P. A. Toledo-Delgado, V. Muñoz-Cruz, and P. V. Torres-Carrion, "Serious games for rehabilitation: Gestural interaction in personalized gamified exercises through a recommender system," *Journal of biomedical informatics*, vol. 97, p. 103266, 2019.

[8] G. Postolache, F. Carry, F. Lourenço, D. Ferreira, R. Oliveira, P. S. Girão, and O. Postolache, "Serious games based on kinect and leap motion controller for upper limbs physical rehabilitation," in *Modern sensing technologies*. Springer, 2018, pp. 147–177.

[9] S. S. Esfahlani, T. Thompson, A. D. Parsa, I. Brown, and S. Cirstea, "Rehabgame: A non-immersive virtual reality rehabilitation system with applications in neuroscience," *Heliyon*, vol. 4, no. 2, 2018.

[10] S. S. Esfahlani, H. Shirvani, J. Butt, I. Mirzaee, and K. S. Esfahlani, "Machine learning role in clinical decision-making: Neuro-rehabilitation video game," *Expert Systems with Applications*, vol. 201, p. 117165, 2022.

[11] S. S. Esfahlani, J. Butt, and H. Shirvani, "Fusion of artificial intelligence in neuro-rehabilitation video games," *IEEE Access*, vol. 7, pp. 102 617–102 627, 2019.

[12] J. F. Ambros-Antemate, M. D. P. Beristain-Colorado, M. Vargas-Trevino, J. Gutierrez-Gutierrez, P. A. Hernández-Cruz, I. B. Gallegos-Velasco, and A. Moreno-Rodriguez, "Software engineering frameworks used for serious games development in physical rehabilitation: systematic review," *JMIR Serious Games*, vol. 9, no. 4, p. e25831, 2021.

[13] S. C. Cramer, L. Dodakian, V. Le, J. See, R. Augsburger, A. McKenzie, R. J. Zhou, N. L. Chiu, J. Heckhausen, J. M. Cassidy *et al.*, "Efficacy of home-based telerehabilitation vs in-clinic therapy for adults after stroke: a randomized clinical trial," *JAMA neurology*, vol. 76, no. 9, pp. 1079–1087, 2019.

[14] W. H. Organization *et al.*, "Who guideline: recommendations on digital interventions for health system strengthening: web supplement 2: summary of findings and grade tables," World Health Organization, Tech. Rep., 2019.

[15] K. E. Laver, Z. Adey-Wakeling, M. Crotty, N. A. Lannin, S. George, and C. Sherrington, "Telerehabilitation services for stroke," *Cochrane Database of Systematic Reviews*, no. 1, 2020.

[16] A. Zeng, X. Xu, L. Zhang, and Q. Li, "Smoothnet: A plug-and-play network for refining human poses in videos," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022.

[17] A. Kanazawa, M. J. Black, D. W. Jacobs, and J. Malik, "End-to-end recovery of human shape and pose," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7122–7131.

[18] M. F. Levin, J. A. Kleim, and S. L. Wolf, "What do motor "recovery" and "compensation" mean in patients following stroke?" *Neurorehabilitation and neural repair*, vol. 23, no. 4, pp. 313–319, 2009.

[19] M. Cirstea and M. F. Levin, "Compensatory strategies for reaching in stroke," *Brain*, vol. 123, no. 5, pp. 940–953, 2000.

[20] S. K. Subramanian, C. L. Massie, M. P. Malcolm, and M. F. Levin, "Does provision of extrinsic feedback result in improved motor learning in the upper limb poststroke? a systematic review of the evidence," *Neurorehabilitation and neural repair*, vol. 24, no. 2, pp. 113–124, 2010.

[21] N. Kolotouros, G. Pavlakos, M. J. Black, and K. Daniilidis, "Learning to reconstruct 3d human pose and shape via model-fitting in the loop," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 2252–2261.

[22] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, "Smpl: A skinned multi-person linear model," in *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, 2023, pp. 851–866.

[23] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee *et al.*, "Mediapipe: A framework for building perception pipelines," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2019.

[24] J. A. Silva, M. F. Silva, H. P. Oliveira, and C. D. Rocha, "Developing a serious video game to engage the upper limb post-stroke rehabilitation," *Applied Sciences*, vol. 15, no. 15, p. 8240, 2025.

[25] Y. H. Ang, C. K. Chan, S. C. Yap, C. K. Toa, P. Tran, and S. K. Goh, "Markerless human motion analysis for telerehabilitation: A case study on squat," in *International Conference on Future Smart Cities*. Springer, 2022, pp. 249–259.

[26] A. Mohamed, H. Abdelmoreed, M. Ehab, Y. Shawky, M. Hadhoud, and A. Al-Kabbany, "Toshfa: A mobile vr-based system for pose-guided exercise rehabilitation for low back pain," *arXiv preprint arXiv:2601.17553*, 2026.

[27] J. E. Fontenele, S. Teixeira, D. L. Sousa, R. Moreira, M. G. Silva, S. Nascimento, L. Pereira, and A. S. Teles, "Rehabilite game+: Pose estimation-based serious games for forearm and hand movement rehabilitation," *Entertainment Computing*, p. 100992, 2025.

[28] M. Maier, H. J. A. Van Hedel *et al.*, "A personalized home-based rehabilitation program using a kinect and tele-rehabilitation: The virtele study," *JMIR Serious Games*, vol. 9, no. 3, p. e26153, 2021.

[29] Y. Tong, Y. Ding, Z. Han, H. Duan, and X. Geng, "Optimal rehabilitation strategies for early postacute stroke recovery: An ongoing inquiry," *Brain circulation*, vol. 9, no. 4, pp. 201–204, 2023.

[30] M. Hank, P. Miratsky, K. R. Ford, C. Clarup, O. Imal, F. F. Verbruggen, F. Zahalka, and T. Maly, "Exploring the interplay of trunk and shoulder rotation strength: a cross-sport analysis," *Frontiers in physiology*, vol. 15, p. 1371134, 2024.