PACE-CBR: A Pose-Aware Conversational Case-Based Exercise Coach

Pedram Salimi[1], Nirmalie Wiratunga[1], Ikechukwu Nkisi-Orji[1], Truong Dang[1], Vihanga Wijayasekara[1] and Kay Cooper[1]

[1]*School of Computing, Engineering and Technology, Robert Gordon University, Aberdeen, UK*

Abstract

Exercise self-management is important across many health and rehabilitation contexts, including long-term physiotherapy, chronic disease management, and home-based wellbeing support. In these settings, people often need to perform exercises without continuous supervision, making timely feedback and longitudinal progress monitoring important. Existing pose-estimation systems can support immediate form correction, while large language models can deliver feedback conversationally; however, both are limited when each exercise session is treated as an isolated event. In this paper, we present PACE-CBR, a pose-aware conversational exercise coach that combines YOLO-based pose tracking, nearest-neighbour joint-angle comparison, tool-augmented LLM feedback, and case-based long-term memory. During a session, deviations between the user's movement and an instructor reference are summarised as joint-level performance metrics. After the session, the agent retrieves prior session cases, generates feedback that is underpinned by these prior session cases, and retains the current session together with the delivered feedback as a new case. The prototype demonstrates how single-session pose feedback can be transformed into reusable longitudinal knowledge for CBR-grounded exercise self-management.

Keywords

Case-based reasoning, Conversational AI, Exercise coaching, Pose estimation, Digital health, Self-management

1. Introduction

Exercise self-management requires people to perform movements repeatedly, interpret their progress, and adjust their behaviour over time. This is important across many health and wellbeing contexts, including physiotherapy, rehabilitation, chronic disease self-management, and home-based maintenance of mobility and function. However, continuous supervision from physiotherapists, exercise specialists, or rehabilitation professionals is rarely available in everyday settings. Digital exercise-coaching systems therefore offer an opportunity to support self-management, provided that their feedback is personalised, understandable, and sensitive to the user's physical limitations.

Computer-vision-based exercise systems can provide immediate feedback by comparing a user's movement with a reference performance. Recent work has explored pose-estimation approaches for real-time exercise monitoring, posture correction, and physiotherapy-style movement assessment [1]. In parallel, large language models (LLMs) are increasingly being investigated for physical activity coaching and exercise-related conversational support [2]. These two strands are complementary: pose-estimation systems can produce quantitative movement evidence, while LLMs can translate such evidence into natural, personalised feedback. However, both approaches are limited when each exercise session is treated as an isolated event. A user may repeatedly struggle with the same movement, gradually improve despite remaining below an ideal reference performance, or perform differently because of fatigue, pain, reduced mobility, or a temporary change in health. Such patterns require longitudinal reasoning over performance trends rather than one-off correction.

Case-based reasoning (CBR) provides a natural framework for this problem by supporting persistent case memory, structured representation of exercise sessions, and similarity-based matching between current and previous performance. In health self-management, CBR has been used to tailor recommendations and support personalised advice, for example in the selfBACK project for low-back pain [3, 4],

ICCBR Workshop on Memory-Based Reasoning for Responsible GenAI at ICCBR2026, August 15, 2026, Bremen, Germany
p.salimi1@rgu.ac.uk (P. Salimi); n.wiratunga@rgu.ac.uk (N. Wiratunga); i.nkisi-orji@rgu.ac.uk (I. Nkisi-Orji); t.dang1@rgu.ac.uk (T. Dang); v.wijayasekara@rgu.ac.uk (V. Wijayasekara); k.cooper@rgu.ac.uk (K. Cooper)

CBR was used to recommend self-management plans comprising of physical activity, and exercises [5]. Instead of relying on self-reported exercise adherence, here we are interested to explore how digital technology can be harnessed to enable exercise performance monitoring. Inspired by this line of work, we investigate how pose-derived exercise evidence can be represented as reusable cases for real-time exercise coaching. Rather than using pose deviations only for immediate feedback, our system retains each completed session as part of the user's longitudinal exercise history, enabling future feedback to be tailored to the user's prior performance, progress, and recurring difficulties.

In this paper, we present PACE-CBR, a conversational AI exercise coach prototype for pose-aware, case-based exercise self-management. The system combines YOLO-based pose tracking, nearest-neighbour joint-angle comparison, a tool-augmented LLM coaching agent, speech-based interaction, and long-term case memory. During an exercise session, the system compares the user's movement with an instructor reference video and summarises performance through joint-level deviation metrics. After the session, the agent retrieves the user's profile and recent prior session cases, generates context-aware feedback, and retains the current session together with the delivered feedback as a new case.

The central research question addressed in this work is: how can real-time deviations between a reference exercise and a user's movement be transformed into reusable cases for longitudinal exercise self-management? Our aim is not to replace professional supervision, but to explore how CBR can make automated exercise feedback longitudinal, personalised, and auditable, with each new session interpreted in relation to the user's previous performance.

This work is intended as a prototype and system-level contribution rather than a claim of novelty for each individual component. Pose estimation, conversational agents, and CBR are established techniques; the contribution of PACE-CBR lies in operationalising their combination for longitudinal exercise self-management. In particular, the paper focuses on how pose-derived movement evidence can be converted into structured exercise-session cases that can be retrieved, interpreted, and retained by a conversational coaching agent. The current prototype therefore demonstrates feasibility and a reusable architecture, while formal evaluation of coaching effectiveness, adherence, and clinical utility remains future work.

The contributions of this paper are:

- A pose-to-case pipeline that converts reference-based movement deviations into structured joint-level exercise-session cases.
- A conversational AI coaching architecture that uses tool calls to retrieve session metrics, user context, and prior cases before generating context-aware feedback.
- A CBR-inspired longitudinal memory mechanism that retains each session together with the feedback delivered, enabling future coaching to be grounded in the user's own exercise history.

Through this prototype, we demonstrate how single-session pose feedback can be transformed into longitudinal knowledge for future self-management support. We later discuss the relevance of this methodology to a cancer-survivorship workshop context, where participants emphasised pacing-aware, adaptive, and trustworthy exercise feedback.

All code and supporting materials are available in the project repository.[1]

2. Related Work

This work builds on research in case-based reasoning for health self-management, computer-vision-based exercise feedback, and conversational AI coaching. CBR is a well-established paradigm in which new problems are solved by retrieving, reusing, revising, and retaining previous cases [6]. In healthcare, CBR has been widely explored for diagnosis, treatment support, monitoring, and personalised recommendation, because it naturally supports reasoning from prior patient experiences and adapting them to new contexts [7, 8]. This makes CBR particularly suitable for self-management settings, where users require repeated recommendations over time rather than isolated one-off advice.

[1] https://github.com/pedramsalimi/PACECBR

A closely related example is the selfBACK project, which developed a case-based decision-support system for low-back-pain self-management. selfBACK used CBR to generate tailored weekly self-management plans from questionnaire data, weekly reports, symptom progression [3]. The project also investigated activity recognition from wearable sensors to monitor daily activity type, activity level, and step counts, thereby linking sensor-derived evidence to personalised self-management support [5]. selfBACK was later evaluated as an app-delivered intervention for low-back-pain-related disability [4]. Although exercises were a central aspect of the self-management plan the mobile based app relied on the patient to accurately self-report on exercise adherence. Our work is inspired by this CBR-based self-management model, but shifts the monitoring to exercise-form performance. Each exercise session is represented as a case containing joint-level pose deviations, session context, and the feedback delivered to the user, allowing future sessions to be interpreted against the user's own prior performance.

CBR has also been applied in sport and physical-activity contexts. Previous work has used CBR to support marathon race planning and to recommend tailored training plans from prior running histories [9, 10]. These systems demonstrate the value of representing previous performance episodes as reusable cases for future planning. However, their focus is athletic performance optimisation rather than rehabilitation-sensitive exercise self-management. In contrast, our system uses cases to support longitudinal, context-aware feedback for self-managed exercise, where progress may be gradual, variable, and shaped by individual physical limitations, fatigue, pain, or changing health status.

Exercise and physical activity are central to many forms of health self-management, including rehabilitation, long-term condition management, and maintenance of functional ability. In home and community settings, people may be asked to perform therapeutic or strengthening exercises without continuous professional supervision. This creates a need for systems that can provide feedback on movement quality, help users understand progress, and support adherence over time. Digital coaching is therefore valuable not only as a motivational tool, but also as a way of helping users execute exercises more safely and consistently when a physiotherapist or exercise specialist is not present.

Computer vision has been widely investigated for exercise and rehabilitation assessment. Markerless pose-estimation systems can extract body landmarks from video and compare a user's movement with a reference performance, enabling automated assessment of exercise form [11]. Recent systems use pose-estimation models, including YOLO-based approaches, to support real-time posture correction and personalised fitness feedback [1, 12]. Such systems show that keypoint-based movement comparison can provide useful immediate feedback. However, most pose-feedback systems are session-local: detected errors are used for real-time correction, but are not usually retained as structured cases for future retrieval, comparison, and longitudinal explanation. Our work therefore extends pose-based feedback by converting movement deviations into reusable cases for CBR.

Conversational agents and large language models offer a further opportunity to make exercise feedback more accessible and personalised. Prior work on conversational agents for physical activity suggests that they can support engagement and behaviour change, but also highlights limitations around repetitive content, attrition, safety, and grounding [13]. Recent LLM-based physical-activity coaching systems show promise for generating personalised coaching dialogue, but still require stronger grounding in reliable user data and expert-informed constraints [2, 14]. Closely related work on modular LLM dialogue systems has argued for ontology-guided intent disambiguation, specialist-agent routing, short- and long-term memory, and CBR-supported retrieval in health-oriented conversational systems [15]. Adjacent work on agentic CBR and explanation generation also shows how supervisor agents, user-feedback loops, feature-actionability constraints, and template-based natural language generation can reduce the risk of unconstrained or infeasible LLM advice [16, 17]. In our system, the LLM is therefore not used as an unconstrained coach. Instead, it acts as an explanation and interaction layer over structured pose-analysis outputs and a longitudinal case base. The agent retrieves the current session summary, compares it with prior cases, generates context-aware feedback, and stores the delivered feedback for future reuse.

Overall, prior work shows that CBR can support longitudinal health self-management, pose estimation can support automated exercise assessment, and conversational AI can deliver personalised coaching. The gap is that these strands are rarely integrated. Our contribution is to represent pose-derived exercise

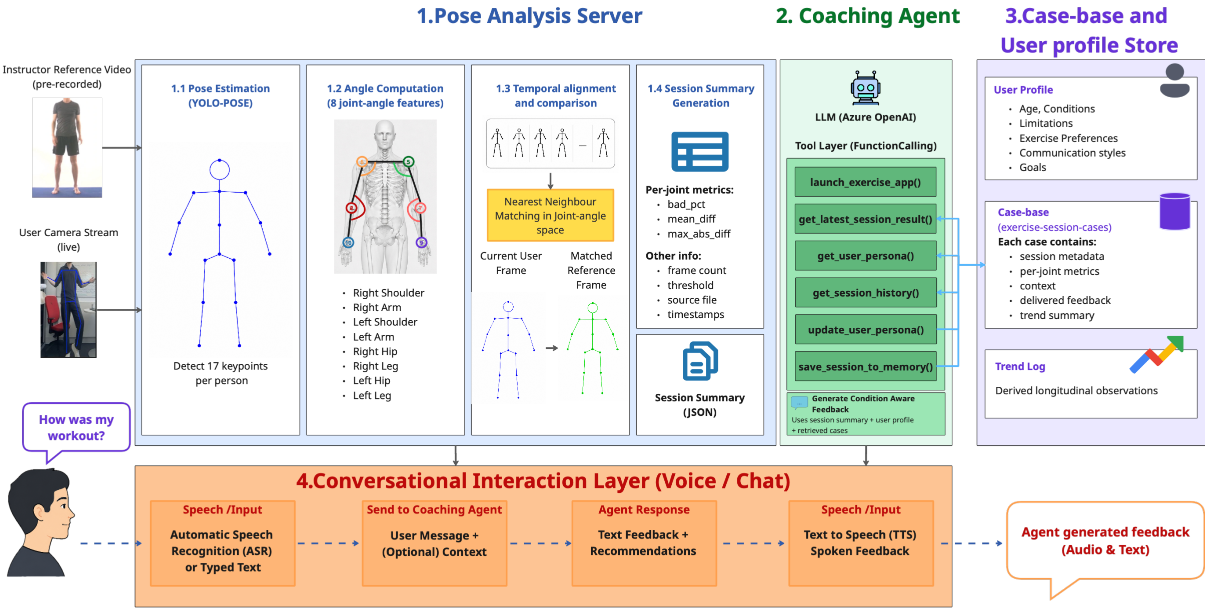

Figure 1: System architecture of PACE-CBR. The pose-analysis server processes an instructor reference video and the user's live camera stream to estimate body keypoints, compute eight joint-angle features, perform nearest-neighbour temporal alignment, and generate a structured session summary. The coaching agent uses tool calls to retrieve the user profile, recent exercise-session cases, and the current session summary before generating context-aware feedback. The delivered feedback and session metrics are then stored as a new case in the case-base, while the conversational layer manages speech or text input and spoken feedback output.

performance as reusable CBR cases, enabling an agentic exercise coach to provide feedback that is not only immediate and conversational, but also grounded in the user's own longitudinal exercise history.

3. Pose Estimation

The proposed system is implemented as a conversational AI exercise-coaching prototype for pose-aware exercise self-management. The system is designed for home or community settings where users may perform physiotherapy-style or wellbeing exercises without continuous professional supervision. It integrates pose-based exercise analysis, tool-augmented language-model coaching, speech-based interaction, and a longitudinal case-base. Its purpose is to transform an exercise session from a one-off movement-comparison event into a reusable case that can support future self-management feedback.

Figure 1 shows the overall architecture of PACE-CBR. The left side of the pipeline processes the instructor reference video and the user's live camera stream. The pose-analysis server extracts body keypoints, converts them into eight joint-angle features, aligns each user frame with the most similar instructor frame, and produces a session-level summary. The centre of the architecture contains the coaching agent, which uses tools to access sensor data and to complete other tasks. Through these tools, the agent retrieves the current session summary, the user's stored profile, and recent prior cases. The right side shows the long-term memory layer, where user profile information, exercise-session cases, and trend observations are retained. The lower part of the figure shows the conversational interaction layer, which supports speech or text input and returns the agent's feedback as spoken output.

This modular design separates four responsibilities. The pose-analysis server is responsible for numerical movement comparison; the coaching agent is responsible for interpretation and feedback generation; the case-base is responsible for longitudinal memory; and the interaction layer is responsible for user-facing conversation. This separation allows the system to ground language-model feedback in structured pose evidence and prior cases, while keeping the sensing, reasoning, storage, and interaction components independently extensible.

Table 1
Eight joint angles derived from COCO keypoints. The angle is computed at vertex B^*, between the vectors $B^*\rightarrow A$ and $B^*\rightarrow C$.

Body region	Joint	A	B^* (vertex)	C
Upper body	right_shoulder	5 L. Shoulder	**6 R. Shoulder**	8 R. Elbow
	right_arm	6 R. Shoulder	**8 R. Elbow**	10 R. Wrist
	left_shoulder	7 L. Elbow	**5 L. Shoulder**	6 R. Shoulder
	left_arm	9 L. Wrist	**7 L. Elbow**	5 L. Shoulder
Lower body	right_hip	11 L. Hip	**12 R. Hip**	14 R. Knee
	right_leg	12 R. Hip	**14 R. Knee**	16 R. Ankle
	left_hip	13 L. Knee	**11 L. Hip**	12 R. Hip
	left_leg	11 L. Hip	**13 L. Knee**	15 L. Ankle

3.1. Pose-Based Exercise Analysis

The pose-analysis module performs a keypoint-based movement comparison between an instructor reference performance and the user's live movement.

In the current prototype, pose estimation is performed using a pre-trained YOLO pose model over video frames. YOLO, short for "You Only Look Once", refers to a family of real-time object detection models that have also been extended to human pose estimation by predicting body keypoints from images or video frames [18].

For each detected person, the model extracts 17 COCO-style body landmarks. These landmarks are then converted into eight planar joint-angle features used by the system: `right_shoulder`, `right_arm`, `left_shoulder`, `left_arm`, `right_hip`, `right_leg`, `left_hip`, and `left_leg`. Each feature is computed from a triplet of landmarks (a, b, c), where b is the angle vertex as shown in Table 1. For example, `left_hip` is computed from keypoints $(13, 11, 12)$, so the vertex is keypoint 11, corresponding to the left hip.

The eight-angle representation was selected as a deliberately compact and interpretable representation for the first prototype. It captures bilateral upper- and lower-body movement using landmarks that are available from the COCO keypoint format and that can be described meaningfully in feedback to users and clinicians. The aim was not to learn an optimal biomechanical representation, but to retain a small set of pose-derived features that are sufficiently transparent to support case representation, longitudinal comparison, and conversational explanation. Similarly, the 25-degree threshold used for bad_pct is an engineering default used to identify substantial deviations in the prototype. It should not be interpreted as a clinically validated boundary between correct and incorrect movement. Future work should validate both the selected features and the threshold values against expert physiotherapist ratings and exercise-specific movement criteria.

Figure 2 summarises the two main operations in the pose-analysis module. First, each detected pose is converted into a vector of eight joint-angle features. For a given joint-angle feature j, the angular deviation between the user and reference pose is computed as:

$$\Delta\theta_j = \left|\theta_{\text{user},j} - \theta_{\text{ref},j}\right|,$$

where $\theta_{\text{user},j}$ is the user's angle for feature j and $\theta_{\text{ref},j}$ is the corresponding angle from the instructor reference frame.

The instructor video is processed before the exercise session to obtain a reference sequence of joint-angle vectors. During the user's session, incoming frames are analysed online and compared with this reference sequence. Temporal alignment is handled through nearest-neighbour matching in joint-angle space. For each user frame at time t, the system selects the instructor frame r^* that minimises the

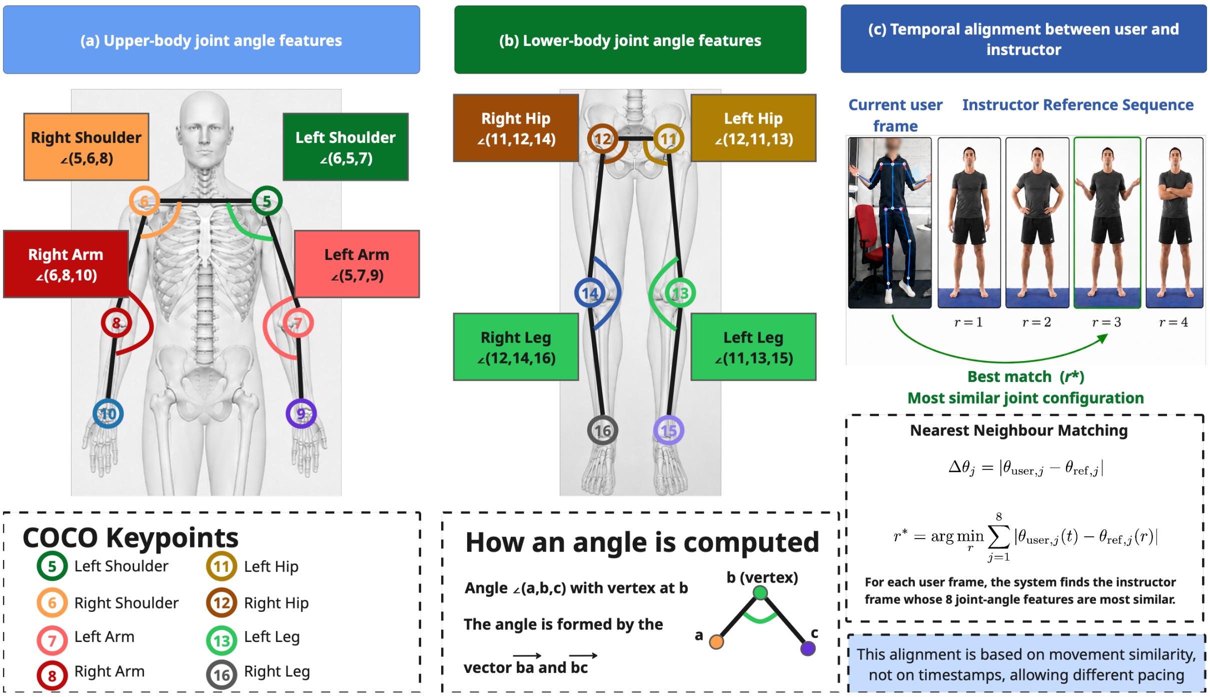

Figure 2: Pose-based joint-angle comparison and temporal alignment in PACE-CBR. Panels (a) and (b) illustrate the eight joint-angle features extracted from COCO body keypoints for upper-body and lower-body pose comparison. Panel (c) illustrates temporal alignment: each current user frame is matched to the instructor frame with the most similar joint-angle configuration, rather than being compared by timestamp.

summed absolute angular difference across all eight joint-angle features:

$$r^* = \arg\min_r \sum_{j=1}^{8} \left|\theta_{\text{user},j}(t) - \theta_{\text{ref},j}(r)\right| .$$

This means that the user is compared with the most similar instructor pose, rather than with the instructor frame occurring at the same timestamp. In the current implementation, the instructor reference video is processed offline by running pose estimation on every frame at the video's native frame rate. In our demonstration setting, this corresponds to 36 fps for the instructor video. When a recorded user video is processed offline, pose estimation is likewise applied at the video's native frame rate, for example 30 fps. During real-time streaming, however, the system sends user frames at a configurable rate: the default is 15 fps, and the user can adjust this between 3 and 30 fps depending on GPU load and desired responsiveness. This design avoids requiring strict synchronisation between the user and the instructor. It is particularly important in home-based exercise and rehabilitation contexts, where users may move more slowly, pause, or perform movements through a reduced range of motion. The system also performs bilateral mirroring detection, allowing comparison when the user faces the opposite direction to the instructor. When mirroring is selected, left and right joint-angle labels are swapped before comparison, so the selected user angle vector is aligned with the instructor's frame of reference.

At the end of a session, the system summarises performance independently for each joint-angle feature. Three metrics are retained: `bad_pct`, the percentage of sampled frames in which the absolute angular difference exceeds a threshold; `mean_diff`, the mean signed angular difference; and `max_abs_diff`, the maximum absolute angular difference observed. The default threshold is 25 degrees. For each joint-angle feature j, the main session-level quality measure is:

$$\texttt{bad\_pct}_j = \frac{\left|\left\{f : \left|\theta^f_{\text{user},j} - \theta^f_{\text{ref},j}\right| \geq \tau\right\}\right|}{N} \times 100,$$

where τ is the angular threshold and N is the number of sampled frames. The `bad_pct` metric is treated as the main indicator of movement quality because it captures how consistently a joint-angle feature deviated from the reference movement. The signed `mean_diff` provides information about the direction of the deviation, while `max_abs_diff` is retained mainly for diagnostic purposes, since very large single-frame deviations may reflect pose-estimation artefacts. Although the pose server records an overall all-joints-within-threshold score, this value is deliberately excluded from the coaching agent's feedback logic because it requires all eight joint-angle features to fall within threshold simultaneously and is therefore too strict for users with restricted mobility.

4. Memory-Based Coaching

The coaching component is implemented as a tool-augmented language agent using an Azure OpenAI large language model (GPT-4o mini)[2].

GPT-4o mini was used in the implemented prototype because it provided a low-cost, tool-calling model available through the deployment environment at the time of development. The architecture is not dependent on this specific model: any language model with reliable tool-use and instruction-following capabilities could be substituted. Evaluating newer reasoning-oriented models for feedback quality, safety, and consistency is an important direction for future work.

Rather than passing raw frame-level sensor data directly to the model, the agent interacts with the rest of the system through a small set of tools. These tools allow the agent to launch the exercise application, retrieve the latest session summary, read or update the user profile, retrieve previous session cases, and save the current session and feedback to long-term memory. This design separates data processing from language generation: the pose-analysis module produces structured metrics, while the language model is responsible for interpretation, explanation, and conversational delivery.

The post-session workflow is explicitly constrained by the agent prompt. When the user reports that an exercise session has finished, the agent must first retrieve the latest joint-level performance summary. Then it retrieves the stored user profile, including relevant medical context and movement limitations, before consulting recent previous sessions for longitudinal comparison. Only after this retrieval process does the agent compose feedback. The generated feedback is then saved together with the session metrics before being delivered to the user. This ordering is important because it ensures that the feedback spoken to the user is the same feedback retained in the case-base, creating an auditable record of the coaching interaction.

4.1. Prompt Design for Context-Aware Feedback

The system prompt encodes both metric-interpretation guidance and population-sensitive feedback rules. It is organised into functional sections covering the current exercise context, the user profile, interpretation of joint-level metrics, the required post-session workflow, feedback style, and persona updates. This structure is used to reduce the risk that the language model misinterprets the numerical output of the pose-analysis module.

The metric-interpretation guidance instructs the agent to treat *bad_pct* as the primary quality signal, to interpret *mean_diff* as a directional tendency, and to treat extreme *max_abs_diff* values cautiously because they may arise from noisy pose estimation. The prompt also warns the agent not to base feedback on the overall all-joints accuracy score. This was introduced to prevent the model from describing otherwise acceptable sessions as poor merely because not all joints were simultaneously within threshold.

The feedback rules are designed for a rehabilitation-sensitive and wellbeing-oriented self-management context. The agent is instructed to compare the user only with their own previous sessions rather than with a healthy-population norm. If a joint corresponds to a recorded condition or limitation, the agent should frame the observation empathetically rather than as a deficit. The prompt also encourages

[2]https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/

Table 2
PACE-CBR exercise-session case representation.

Case component	Stored attributes
Problem/context	User profile, condition notes, exercise name, date
Pose evidence	Per-joint `bad_pct`, `mean_diff`, `max_abs_diff`
Session metadata	Session identifier, source file, frame count, angle threshold
Solution	context-aware feedback generated by the agent
Retained outcome	Delivered feedback and trend observation

acknowledgement of partial movement and effort, especially where the user is working within a restricted range of motion. In this way, the agent's role is not simply to report errors, but to translate movement data into supportive, self-referential feedback.

4.2. Case Representation and Memory

PACE-CBR maintains two forms of memory. A short-term conversational memory maintains dialogue state within the active interaction thread, allowing the agent to respond coherently during the current conversation. A separate long-term memory stores the user's profile and completed exercise sessions. The user profile contains fields such as name, age, gender, condition, and notes. These fields allow the agent to contextualise feedback according to the user's exercise goals, movement limitations, preferences, and relevant health context.

Each completed exercise session is stored as a structured case. A case contains the session identifier, date, day of week, exercise name, source session file, number of analysed frames, angle threshold, joint-level performance statistics, and the exact feedback delivered to the user. The joint-level statistics store *bad_pct*, *mean_diff*, and *max_abs_diff* for each of the eight angle features. Storing the delivered feedback as part of the case is important because it preserves not only what the system measured, but also how the system explained that measurement to the user.

In addition to the case-base, the system maintains an append-only trend log. When a new session is saved, the system compares its joint-level *bad_pct* values with those from the previous session and records the resulting changes. This separates longitudinal trend observations from the main case-base, keeping the stored session cases compact while preserving a readable history of session-to-session changes.

4.3. Case-Based Reasoning Workflow

The current workflow implements a CBR-inspired workflow over exercise-session cases. The retrieved cases are recent prior sessions rather than nearest neighbours in a learned similarity space. This design was chosen for the first prototype because the main goal is to support longitudinal self-comparison over repeated sessions. In CBR terms, the problem component of a case consists of the user context, exercise type, and joint-level pose evidence; the solution component is the feedback generated and delivered by the agent; and the retained outcome is the stored session case together with a trend observation.

This workflow allows the system to move beyond one-off pose correction. A session is not treated as an isolated event; it becomes part of a growing case history that can support future feedback. For example, a high `bad_pct` for one joint may be interpreted differently if the same joint has improved over previous sessions, if the user has reported a relevant movement limitation, or if the previous feedback already focused on that area. The case-base therefore enables self-referential progress monitoring, where the user is compared against their own prior performance rather than against an external ideal.

4.4. Conversational Interaction Layer

The interaction layer provides a spoken and text-based conversational interface around the exercise and coaching pipeline. The user can interact with the system through speech or typed input. Speech

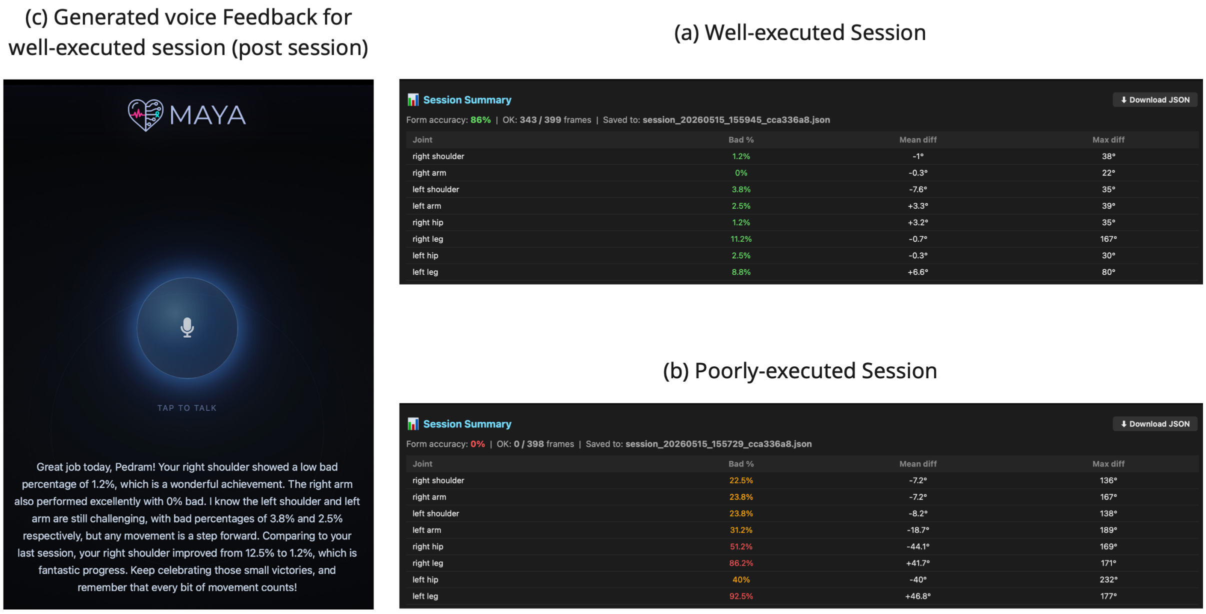

Figure 3: Prototype interface outputs showing generated post-session feedback and joint-level summaries for well-executed and poorly executed exercise sessions.

Table 3
Example longitudinal comparison between two retained exercise-session cases.

Joint feature	Previous bad_pct	Current bad_pct	Change	Interpretation
Right shoulder	35.0	18.0	-17.0	Improved
Left shoulder	22.0	28.0	+6.0	Needs attention
Right hip	12.0	10.0	-2.0	Stable/improved
Left leg	40.0	31.0	-9.0	Improved

input is transcribed using automatic speech recognition, passed to the coaching agent, and the agent's response is then synthesised as spoken output. The interface also supports a tap-to-speak interaction pattern, which is appropriate for shared or home-based health technologies because it avoids always-on listening and gives the user explicit control over when speech is captured.

The conversational layer is deliberately separated from the pose-analysis server. The pose server is responsible for exercise tracking and session-summary generation, while the chat server manages speech recognition, agent invocation, and text-to-speech delivery. This separation keeps the architecture modular: improvements to pose estimation, case retrieval, or dialogue delivery can be made independently. Together, these components provide an implemented prototype of a CBR-grounded, voice-enabled exercise coach for longitudinal physical activity self-management.

5. Stakeholder Consultation, Limitations and Future Work

PACE-CBR is intended as a general methodology for pose-aware, case-based exercise coaching in self-management contexts. For evaluation purposes, we developed a prototype for exercise-based self-management of cardiotoxicity risk in adolescent and young adult cancer survivors.

5.1. Co-design Workshop Feedback

In a co-design workshop linked to the EU Horizon-funded MAYA project, participants valued exercise feedback that was practical, pacing-aware, and able to reflect progress over time, rather than generic reminders or one-off exercise advice. This supports the design direction of PACE-CBR, where feedback is generated from the current exercise performance while also drawing on previous sessions and feedback history.

The workshop findings reported here are used as design evidence rather than as a formal summative evaluation of PACE-CBR. The co-design activity was conducted within the MAYA project with a mixed stakeholder group, including adolescent and young adult cancer-survivorship stakeholders, carers and/or family representatives, health-professional stakeholders, technology representatives, facilitators, and note-takers. Discussion activities focused on requirements for smart-mirror-based health coaching, including exercise support, personalisation, trust, pacing, and progress feedback. Data were captured through facilitator notes and post-workshop synthesis, then analysed thematically to identify design implications relevant to conversational exercise coaching.

The thematic analysis highlighted the need for personalised exercise feedback that accounts for the young person's treatment history, current symptoms, goals, confidence, and functional ability. For PACE-CBR, this means moving beyond isolated movement assessment towards feedback that interprets joint performance in context, including fatigue, pain, reduced mobility, treatment stage, recent illness, or previous surgery. This is important for young cancer survivors, whose exercise performance may fluctuate for reasons not captured by pose data alone. Related to this participants noted that onboarding should establish users' goals, baseline ability, treatment-related limitations, confidence, and feedback preferences. This would support safer personalisation by informing both performance thresholds and the level of explanation provided during or after exercise.

To improve interpretability, a natural extension is to map per-joint bad_pct values into physiotherapist-informed qualitative descriptors, such as excellent, good, needs attention, significant difficulty, or unable to perform. This would make feedback more consistent, reduce reliance on numeric percentage values, and improve readability. Periodic self-efficacy questions could allow PACE-CBR to track changes in confidence alongside movement performance, supporting feedback that reflects biomechanical accuracy as well as confidence, consistency, and independence. Together, these findings motivate future work on physiotherapist-informed descriptors, user-reported context, and memory-based tracking of progress and outliers.

5.2. Limitations and Desirable Technical Extensions

The most important limitation of the current work is the absence of a formal empirical evaluation. The prototype demonstrates the feasibility of converting pose-derived exercise evidence into retained cases and using those cases to support conversational feedback, but it does not yet show that memory-augmented feedback improves user outcomes. In particular, we have not yet compared feedback generated with and without case memory, feedback grounded only in the current session versus feedback grounded in prior sessions, or recency-based retrieval versus similarity-based retrieval. Future evaluation should therefore include ablation studies of the memory component, expert assessment of feedback quality, and user-centred measures such as perceived personalisation, trustworthiness, usefulness, explainability, and motivation to continue exercising.

The second limitation of the current prototype is that case retrieval is recency-based rather than similarity-based. The agent retrieves recent prior sessions and compares the current joint-level metrics with previous cases. This is useful for longitudinal self-comparison, but it does not yet implement a full similarity-based retrieval mechanism. Future work will define similarity measures over exercise-session cases, using features such as exercise type, joint-level `bad_pct`, `mean_diff`, `max_abs_diff`, user-reported difficulty, and relevant movement limitations. This would allow the system to retrieve not only the most recent sessions, but also previous sessions that are most similar to the current performance pattern.

Another direction is the development of a qualitative interpretation taxonomy for joint-level performance metrics. At present, PACE-CBR reports per-joint `bad_pct` values as raw percentages, which the LLM must interpret during feedback generation. A principled mapping, for example 0–10% as excellent, 11–25% as good, 26–45% as needs attention, 46–70% as significant difficulty, and greater than 70% as unable to perform, could make feedback more consistent and easier to interpret. However, such thresholds should not be chosen arbitrarily. They should be co-designed with physiotherapists, rehabilitation specialists, and target users, and validated against expert ratings of real exercise sessions.

The current pose-analysis module also has technical limitations. It relies on 2D pose estimation from camera frames, which may be affected by occlusion, camera angle, lighting, clothing, and noisy keypoint detection. In addition, the current implementation processes every frame of the instructor reference video during offline preparation. This gives a dense reference sequence for nearest-neighbour matching, but it can be computationally expensive for longer videos or higher frame rates. Future work should investigate frame-striding, key-frame selection, or pose-clustering methods to reduce the size of the reference sequence while preserving sufficient movement coverage for accurate temporal alignment. During live use, the streaming rate is already configurable, but further optimisation is needed to balance latency, responsiveness, GPU load, and feedback quality.

6. Conclusion

This paper presented PACE-CBR, a conversational AI exercise-coaching prototype that explores how pose-derived exercise feedback can be transformed into reusable cases for longitudinal self-management. The system combines YOLO-based pose tracking, nearest-neighbour joint-angle comparison, tool-augmented LLM feedback, and case-based long-term memory. Rather than treating each exercise session as an isolated event, the prototype stores joint-level performance metrics together with the feedback delivered to the user, allowing future coaching interactions to be grounded in the user's own exercise history.

The main contribution of this work is a pose-to-case pipeline for exercise self-management. Each completed session is represented as a structured case containing the exercise context, per-joint deviation metrics, and the agent's context-aware feedback. The conversational agent retrieves recent prior cases, compares them with the current session, and uses this longitudinal context to generate feedback that is self-referential rather than norm-comparative. This is relevant across home-based physiotherapy, rehabilitation, chronic disease self-management, and wellbeing contexts, where progress may be gradual, variable, and shaped by individual physical limitations or changing health status.

While the current prototype remains limited by recency-based retrieval, unvalidated interpretation thresholds, and the technical constraints of 2D pose estimation and frame-wise matching, it demonstrates the feasibility of integrating computer vision, conversational AI, and CBR-inspired memory in a single exercise-coaching workflow. Overall, this work argues that CBR offers a useful framework for making automated exercise feedback more continuous, auditable, and personalised. By retaining exercise sessions as reusable cases, conversational exercise coaches can move beyond immediate correction towards longitudinal self-management support.

Acknowledgments

This work was funded by the European Union's Horizon Europe programme under grant agreement No. 101213323 (MAYA). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HaDEA). Neither the European Union nor the granting authority can be held responsible for them.

Declaration on Generative AI

During the preparation of this work, the authors used GPT-5 for grammar and spelling checks. The authors also used GPT Image 2 to generate selected visual elements used in Figures 1 and 2, such as skeleton illustrations. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] H. Kotte, M. Kravcik, N. Duong-Trung, Real-time posture correction in gym exercises: A computer vision-based approach for performance analysis, error classification and feedback., in: MILeS@ EC-TEL, 2023, pp. 64–70.

[2] X. Lai, J. Chen, Y. Lai, S. Huang, Y. Cai, Z. Sun, X. Wang, K. Pan, Q. Gao, C. Huang, Using large language models to enhance exercise recommendations and physical activity in clinical and healthy populations: scoping review, JMIR Medical Informatics 13 (2025) e59309.

[3] P. J. Mork, K. Bach, A decision support system to enhance self-management of low back pain: protocol for the selfback project, JMIR research protocols 7 (2018) e9379.

[4] L. Sandal, K. Bach, C. Øverås, M. Svendsen, T. Dalager, J. Stejnicher Drongstrup Jensen, et al., Effectiveness of app-delivered, tailored self-management support for adults with lower back pain-related disability: a selfback randomized clinical trial. jama intern med. 2021; 181 (10): 1288–96, 2021.

[5] S. Sani, N. Wiratunga, S. Massie, K. Cooper, Selfback—activity recognition for self-management of low back pain, in: International Conference on Innovative Techniques and Applications of Artificial Intelligence, Springer, 2016, pp. 281–294.

[6] A. Aamodt, E. Plaza, Case-based reasoning: Foundational issues, methodological variations, and system approaches, AI communications 7 (1994) 39–59.

[7] I. Bichindaritz, C. Marling, Case-based reasoning in the health sciences: What's next?, Artificial intelligence in medicine 36 (2006) 127–135.

[8] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, M. Folke, Case-based reasoning systems in the health sciences: a survey of recent trends and developments, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 41 (2010) 421–434.

[9] B. Smyth, Marathon race planning: A case-based reasoning approach, IJCAI, 2018.

[10] C. Feely, B. Caulfield, A. Lawlor, B. Smyth, Using case-based reasoning to predict marathon performance and recommend tailored training plans, in: International conference on case-based reasoning, Springer, 2020, pp. 67–81.

[11] B. Debnath, M. O'brien, M. Yamaguchi, A. Behera, A review of computer vision-based approaches for physical rehabilitation and assessment, Multimedia Systems 28 (2022) 209–239.

[12] H. Kotte, F. Daiber, M. Kravcik, N. Duong-Trung, Fitsight: Tracking and feedback engine for personalized fitness training, in: Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization, 2024, pp. 223–231.

[13] T. C. Luo, A. Aguilera, C. R. Lyles, C. A. Figueroa, Promoting physical activity through conversational agents: mixed methods systematic review, Journal of Medical Internet Research 23 (2021) e25486.

[14] M. Jörke, S. Sapkota, L. Warkenthien, N. Vainio, P. Schmiedmayer, E. Brunskill, J. A. Landay, Gpt-coach: towards llm-based physical activity coaching, in: Proceedings of the 2025 CHI conference on human factors in computing systems, 2025, pp. 1–46.

[15] N. Wiratunga, V. A. Wijayasekara, I. Nkisi-Orji, P. Salimi, K. Martin, C. Bolaños, icare: Ontology-guided intent routing for multi-agent llm-based dialogue systems, context (including current conversation goals) 5 (2025) 6.

[16] P. Salimi, N. Wiratunga, D. Corsar, A. Wijekoon, Towards feasible counterfactual explanations: A taxonomy guided template-based nlg method, in: ECAI 2023, IOS Press, 2023, pp. 2057–2064.

[17] P. Salimi, N. Wiratunga, D. Corsar, Agentic cbr in action: empowering loan approvals through interactive, counterfactual explanations. (2025).

[18] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.