TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Stefanova, V., Maeyens, E., Brughmans, J., Vargemidis, D., Lebleu, J., Stulens, L., Broeder, S., Gilis, K., Deschodt, M., Vanwanseele, B., Beckwée, D., Gilat, M., Filtjens, B., & Vanrumste, B. (2026). Deep learning-based physical exercise assessment of older adults using single-camera videos. *PLOS Digital Health*, *5*(8), Article e0001576. https://doi.org/10.1371/journal.pdig.0001576

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

PLOS Digital Health

RESEARCH ARTICLE

Deep learning-based physical exercise assessment of older adults using single-camera videos

Vayalet Stefanova[1], Evelien Maeyens[2], Jeroen Brughmans[2], Dimitri Vargemidis[3], Julien Lebleu[4], Leen Stulens[5], Sanne Broeder[5], Karen Gilis[6], Mieke Deschodt[7], Benedicte Vanwanseele[8], David Beckwée[9], Moran Gilat[10], Benjamin Filtjens[11,12☯*], Bart Vanrumste[1☯]

1 Department of Electrical Engineering (ESAT), KU Leuven, Leuven, Belgium, **2** Faculty of Engineering Technology, KU Leuven, Leuven, Belgium, **3** Department of Computer Science, KU Leuven, Leuven, Belgium, **4** moveUP, Brussels, Belgium, **5** In4Care Community vzw, Grimbergen, Belgium, **6** Zorg Leuven, Leuven, Belgium, **7** Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium, **8** Department of Movement Sciences, KU Leuven, Leuven, Belgium, **9** Rehabilitation Research Group, Vrije Universiteit Brussel, Jette, Belgium, **10** Neurorehabilitation Research Group (eNRGy), Department of Rehabilitation Sciences, KU Leuven, Leuven, Belgium, **11** Department of Engineering Systems and Services, Delft University of Technology, Delft, The Netherlands, **12** Institute for Health Systems Science, Delft University of Technology, Delft, The Netherlands

☯ These authors contributed equally to this work.
* b.filtjens@tudelft.nl

Check for updates

OPEN ACCESS

Citation: Stefanova V, Maeyens E, Brughmans J, Vargemidis D, Lebleu J, Stulens L, et al. (2026) Deep learning-based physical exercise assessment of older adults using single-camera videos. PLOS Digit Health 5(8): e0001576. https://doi.org/10.1371/journal.pdig.0001576

Editor: Crina Grosan, King's College London, UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND

Received: October 9, 2025

Accepted: June 26, 2026

Published: August 6, 2026

Copyright: © 2026 Stefanova et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The dataset collected in this study, including data from all seven subjects of normalized 2D pose sequences, joint angles, and exercise labels, is available under controlled data access at the Research

Abstract

Regular physical activity preserves functional independence in older adults, yet care-home residents often miss out because personalized supervision is scarce. Autonomous, technology-supported exercise platforms could deliver such guidance without additional staff time—but only if sessions are automatically monitored for safety and quality. We therefore designed a deep learning (DL) system that (a) recognizes individual exercise types and (b) estimates joint angle trajectories from a standard video recording. These outputs are used to compute objective exercise performance metrics (EPMs) such as duration, repetition count, motion variability, and range of motion. Seven care-home residents (aged between 65–94 years) performed six common rehabilitation exercises in front of a single camera while wearing 17 inertial sensors (Xsens MVN Awinda) that provided ground-truth joint angles. Two-dimensional skeleton poses estimated from the video were fed into a temporal convolutional neural network to recognize the exercises and estimate three-dimensional joint angles. We evaluated exercise segmentation with F1@50 and angle regression with mean per-joint angular error (MPJAE) across nine trunk and lower-limb joints, using leave-one-subject-out cross-validation. Pearson correlations assessed agreement between estimated and ground-truth EPMs. The DL model achieved an F1@50 of 0.92 (± 0.04) for exercise recognition and an MPJAE of 7.7° (± 0.91) for joint angle estimation. The estimated EPMs aligned closely with ground truth, achieving correlation scores of 0.93 (95% CI [0.90, 0.95]) for duration, 0.86 (95% CI [0.80, 0.90]) for

Data Repository of KU Leuven (https://doi.org/10.48804/VNM5WN). Video recordings cannot be shared due to constraints imposed by participant informed consent and ethical approval. All code required to train and evaluate the models, compute exercise performance metrics, and reproduce the figures presented in this study is publicly available at https://github.com/vayalet/Exercise-performance-DL.

Funding: This work was supported by the province Flemish Brabant as part of the smart region Project (https://www.vlaamsbrabant.be/en/node/14377) "AI@WZC" (BF, BV, KG, JL, LS, SB). BV was additionally supported by the Flemish Government under the Flanders AI Research Program (FAIR: https://www.flandersairesearch.be/). VS, BF, and BV were further supported by the strategic basic research project RevalExo (grant number: S001024N) funded by the Research Foundation - Flanders (https://www.fwo.be/) and the research project AidWear funded by the Federal Public Service for Policy and Support (https://bosa.belgium.be/). The remaining authors (EM, JB, DV, MD, BeV, DB, MG) did not receive specific funding for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

repetition count, and between 0.3 and 0.9 for motion variability and range of motion across exercises. The DL algorithm reliably estimates key exercise outcomes from a single video stream. This video-based monitoring pipeline could enable unsupervised, technology-supported exercise assessment in residential care homes while safeguarding session quality and safety. Future work will validate the approach in larger and more diverse cohorts.

Author summary

Staying physically active is essential for older adults to maintain their independence, but many residents in care homes do not receive the individualized exercise supervision they need. We explored how artificial intelligence can help fill this gap. In our study, we developed a computer system that can watch a person exercise through a single video camera and automatically evaluate how well the exercises are performed. Specifically, the system identifies which exercise is being done and estimates how the person's joints are moving, similar to how a physiotherapist might observe. From this information, it calculates key measures such as how long an exercise lasts, how many repetitions are performed, and how consistently and extensively the person moves. We tested this system with care-home residents performing common rehabilitation exercises, and we found that its assessments were very similar to those obtained using a professional motion-capture system with body-worn sensors. This approach could make it possible to monitor exercise quality without needing extra staff, supporting safe and effective physical activity programs in care homes.

Introduction

The aging population is emerging as one of the leading socio-political challenges of the 21st century. According to demographic projections by the United Nations, the proportion of adults aged 65 and above is set to more than double from 9% in 2019 to 16% by 2050 [1]. With age there is a decline in physical and mental capacities, which diminishes independence among older adults [2,3]. Physical inactivity and sedentary behaviour plays a major role in this decline, whereas in most cases regular exercise has been shown to help delay or prevent these negative effects [4,5]. Regular physical exercise has also been associated with improved quality of life as it promotes healthy aging and disease prevention both for people living at home and in residential care facilities [6,7].

However, maintaining adherence to prescribed exercise regimens remains a major challenge [8]. Healthcare providers often prescribe self-directed exercises as part of rehabilitation therapy, but patient adherence is frequently low, undermining the effectiveness of these interventions [9]. This issue is especially prevalent among older adults living in residential long-term care facilities, who typically do not engage

in sufficient physical activity to meet health recommendations [8]. Caregivers try to encourage older adults to perform physical exercises through personal guidance [10]. Access to such guidance is, however, not universally available due to the decreasing number of caregivers, the growing number of residents, and high financial cost of personalized interventions [11].

To address this challenge, various technologies have been developed to promote physical activity. These include exergames, which combine physical exercise with game-like tasks to enhance motivation, enjoyment, and engagement [12,13]. Virtual reality simulators [14] and dance programs [15,16] offer a similar solution by introducing variety into routine exercise programs, potentially improving adherence and overall satisfaction. Humanoid robot coaches have also been evaluated in terms of their acceptance and ability to improve engagement by providing real-time feedback of upper body exercise performance [17–19]. The envisioned future of humanoid robotic exercise coaching systems is to autonomously track users' 3D articulated movements, deliver personalized motivational and corrective feedback, and continuously monitor safety [20].

Beyond motivating older adults to move more, an important application of these technologies is to support caregivers in evaluating how well individuals adhere to prescribed exercise routines and how effective those routines are in achieving therapeutic goals [21,22]. This aligns with the broader goal of monitoring key biomarkers of healthy aging, which, as suggested by the World Health Organization, can provide insights into an individual's functional status and risk factors for decline [23]. Traditional monitoring methods often rely on specialized equipment such as inertial measurement units (IMUs), depth cameras, or multi-camera setups. These systems are sometimes used in combination, such as pairing IMUs with video cameras, to improve motion tracking, and may also be supplemented with additional physiological sensors for monitoring parameters like heart rate, muscle fatigue, or respiration [21,24–26]. While such setups can yield rich multimodal data for assessing physical performance, they are typically impractical for continual use by older adults due to their cumbersome setup nature and the frequent need for caregiver assistance.

Standard single-camera video offers an accessible contactless solution for exercise monitoring which can be deployed on a range of platforms, including robotic coaches, as well as more affordable options such as mobile phones or simple webcams. Advances in computational methods, particularly deep learning [27], combined with extensive publicly available datasets [28], have enabled the development of pose estimation algorithms such as OpenPose and AlphaPose [29,30]. These algorithms estimate 2D image-plane positions of anatomical points (e.g., ankles and knees) in each video frame. In the field of computer vision, numerous methods have been proposed for monitoring physical activity based on 2D poses which exclude the depth dimension as input [31–35]. However, these algorithms are typically trained and validated on young participants and may exhibit biases, particularly when applied in the context of exercise monitoring of older adults who typically move slower and with less range of motion [36–38]. Existing video pose-based methods for older adults largely target isolated rehabilitation tasks, such as sit-to-stand (STS) [39,40] and gait analysis [41,42]. The clinical outcome metrics that are computed are then limited to the rehabilitation goals of the particular task such as STS phases and their duration [39], cadence, and step time [42]. Other approaches provide a single metric such as a mobility score for a standardized mobility test [43] or movement similarity score with a reference coach during exercise performance [44].

Importantly, existing methods do not offer comprehensive performance monitoring across a wide range of exercises involving both upper and lower body movements in older adults. To address this, we propose a deep learning (DL) framework capable of estimating a clinically meaningful set of exercise performance metrics (EPMs) to support accessible, longitudinal monitoring and the potential for personalizing exercise programs by caregivers. The EPMs that we selected to compute align with the FITT principle (frequency, intensity, time, type) which is a well established fitness guideline [45]. The EPMs include (a) repetition count, which captures how often movements are performed and reflects the frequency component; (b) exercise duration, which measures the time spent engaging in an activity and corresponds to the time component; (c) range of motion, which reflects joint mobility; and (d) motion variability, which evaluates movement quality and stability providing insight into potential fall risk [46]. The combination of motion quality metrics, repetition count and

duration can inform the adjustment of exercise intensity. To enable the computation of these EPMs, our DL framework (a) performs fine-grained exercise detection in terms of frame-level classification labels and (b) estimates joint angles in terms of the 3 rotational degrees of freedom: flexion/extension, lateral and axial bending. This approach supports objective tracking of an older adult's exercise performance without the need for caregiver supervision.

In this work, we address the limitations of existing video-based exercise monitoring approaches by proposing a framework designed for older adults living in residential care facilities. Specifically, our main contributions are threefold. First, we introduce a DL–based framework for fine-grained, exercise detection and joint angle estimation from single-camera video, allowing analysis of body movements. Second, we define and compute a clinically meaningful set of EPMs, that capture not only activity quantity (e.g., repetitions and duration) but also movement quality (e.g., range of motion and variability). Third, we demonstrate how these metrics can support the objective longitudinal monitoring of exercise performance in older adults using accessible technology.

Materials and methods

Data acquisition

Ethics statement. The dataset was collected according to the ethical guidelines and approval procedures defined by the ethical committee of KU Leuven and the university hospital UZ Leuven (EC Research UZ/KU Leuven). Ethical approval was obtained under Clinical Trial Center (CTC) reference number S61015. Prior to the first study-related procedure, all participants signed a written informed consent form.

Subject inclusion criteria. We recruited seven participants (mean age: 82.4 ± 8.1 years, 3 females and 4 males) from a residential care home in Leuven, Belgium. The home's physiotherapist assisted in identifying potential participants, excluding those recovering from recent fractures or other injuries that could impede exercise performance. Participants were included if they had a transferring score of 1 or 2 on the Katz Index of Independence in Activities of Daily Living (ADL), indicating their ability to safely perform exercises without requiring assistance from a caregiver [47]. The transferring score evaluates the individual's ability to move from one place to another, such as from a bed to a chair. A score of 2 indicates that the individual is independent in transferring when using a walking aid and a score of 1 indicates that the individual is fully independent in transferring.

Data collection protocol. Each participant carried out five or six exercise sessions. Each exercise session included six exercises that were deemed safe to carry out unsupervised by a physical therapist. All exercises were selected from the Geriatric Activation Program Pellenberg (GAPP), a rehabilitation program designed for individuals over 65 years old [48]. Exercises were selected to include a wide variety of movements to test the limits of the deep learning framework. Each exercise was performed twice within each exercise session and the order of the exercises was randomized between sessions and participants.

The measurement setup and the six exercises are visualized in Fig 1. The setup included a chair with armrests and sturdy tables positioned close together for support. Exercise equipment, such as a cup, was taped to the ground to prevent slippingTable 1). A webcam (Rifaa WebCam) was placed at a sufficient distance to capture the entire measurement setup. Camera footage was recorded at 30 frames per second with a resolution of 1920 x 1080. Two trained graduate students manually annotated each frame of the videos to identify physical exercises and split them into more fine-grained actions (Table 1) which we refer to as the sub-exercise action labels in this paper. A senior researcher verified the labels for correctness.

Each of the 6 exercises described in Fig 1 is split into sub-exercise actions which are used in the analysis of this study instead of the complete exercises. This segmentation was done by identifying the individual actions that collectively form a complete exercise. The activity labeled as 0 includes all movements unrelated to the instructed exercises.

We collected inertial-based motion capture (MoCap) data using Xsens MVN Analyze Pro (Xsens Technologies B.V., Netherlands) to use as ground truth for our regression task of estimating joint angles [49]. The MoCap system consisted

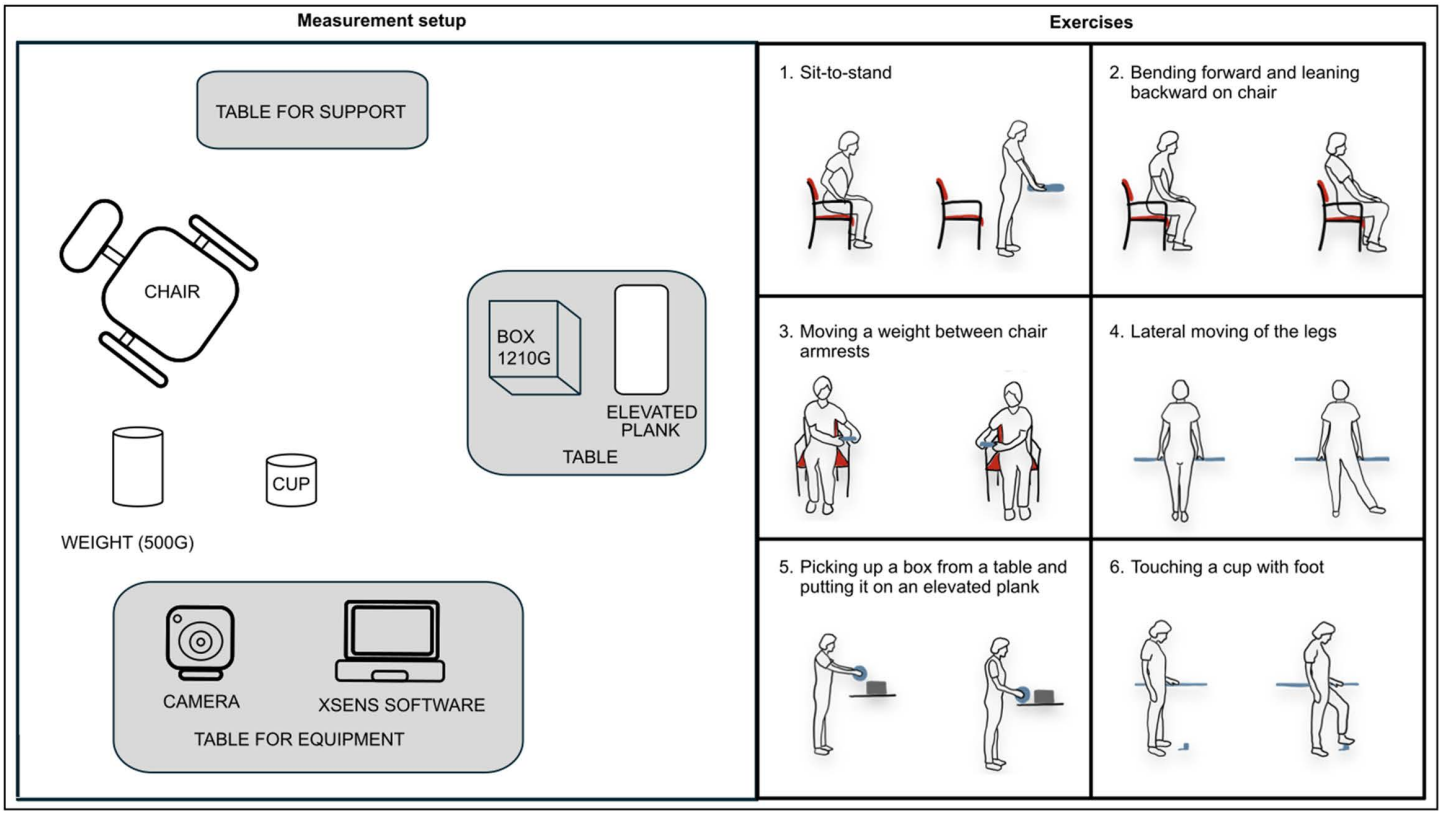

Fig 1. The data collection setup for exercise monitoring. Left: The setup includes a table for support, a main table with an elevated plank and a box (1210 g) for exercises, a table for equipment, which holds a camera and a laptop with a signal receiver and the Xsens MVN Analyze Pro software [49]. A chair with armrests and backrest is positioned for seated exercises, and additional tools like a cylindrical weight (500 g) and a cup are used for specific tasks. **Right:** The 6 physical exercises performed by participants during data collection.

https://doi.org/10.1371/journal.pdig.0001576.g001

of 17 wireless motion sensors (MTw Awinda) operating at 60 Hz, which estimated 3D joint kinematics. These 3D joint kinematics have been validated against gold-standard multi-camera marker-based MoCap systems [50]. Synchronization between the camera footage and motion sensor data was provided by the MVN Analize software through a LAN connection to the webcam. Prior to collecting data, the MTw sensors were calibrated using MVN Analyze Pro software (version 2022.01). This process appeared challenging for the older adults as some participants exhibited hand tremors during the phase requiring them to stand still while others had limited arm swing during the walking phase that differed from the expected healthy adult movement pattern. Consequently, this led to increased noise in the upper body joint kinematics. As a result, certain upper body joints were excluded from the analysis (see the Methods section for details). This highlights the need for caution when using IMU systems that rely on calibration based on movement patterns of healthy adults.

Body keypoints extraction. The video recordings were first processed to estimate sequences of 2D keypoints. The use of 2D skeleton sequences has been shown to be effective for various downstream motion analysis tasks, providing a robust and reliable foundation for modeling motion across diverse datasets [51–53]. This approach has also been shown to work well with older adult populations [54], which motivated our decision to adopt it in this study. To extract 2D keypoints, we analyzed video segments where only the participant was visible, using AlphaPose, a state-of-the-art pose estimation algorithm [30]. AlphaPose provides high spatial resolution by detecting 17 anatomical keypoints for each video frame, including major joints such as the shoulders, elbows, hips, knees, and ankles. Each keypoint is represented

Table 1. Overview of sub-exercise actions and corresponding labels used in the analysis.

Label	Sub-exercise action	Complete exercise
0	Random movements	Random movements
1	Getting seated on chair	Sit-to-stand
2	Standing up from chair	
3	Bending forward while seated on chair	Bending forward/backward on chair
4	Leaning backward while seated on chair	
5	Moving a weight from left to right armrest of chair	Moving a weight between armrests
6	Moving a weight from right to left armrest of chair	
7	Lateral moving of the right leg	Lateral leg movement
8	Lateral moving of the left leg	
9	Picking up a box and placing it on an elevated plank	Picking up a box from table/plank
10	Picking up a box from elevated plank and placing it in front of them	
11	Moving right foot forward to touch a cup	Touching a cup with foot
12	Moving right foot backward after touching a cup	
13	Moving left foot forward to touch a cup	
14	Moving left foot backward after touching a cup	

https://doi.org/10.1371/journal.pdig.0001576.t001

by its *x* and *y* coordinates along with an associated confidence score indicating the reliability of the detection. Although AlphaPose processes frames independently and does not explicitly model temporal dependencies, the resulting frame-by-frame keypoint sequences implicitly capture the temporal progression of motion. This structure provides a frame-by-frame series of 17 $\times$ 3 matrices, ordered according to the COCO whole body keypoint format [55], which serves as input to our deep learning framework. To reduce variability due to camera position and person size, all keypoint poses were normalized using a subject-centered square crop based on visible joints and then scaled to the range [–1, 1]. All data were pseudonymized to ensure participant privacy, and the two modalities (2D pose and joint angles), together with the corresponding exercise labels, are available upon request at https://doi.org/10.48804/VNM5WN.

Deep learning framework

As shown in Fig 2, this study focused on the application of an existing fully convolutional architecture proposed by Pavllo et al. [31] to address two downstream tasks: (1) exercise recognition for frame-by-frame action annotation and (2) joint angle prediction for continuous kinematic analysis. The outputs of these tasks were used to compute the EPMs described in Section Exercise performance metrics (EPMs). The chosen DL model, referred to as VideoPose, is motivated by its state-of-the-art performance in the regression task of 3D human pose estimation [31]. Moreover, it has also been successfully applied to classification tasks, such as detecting freezing of gait in patients with Parkinson's disease and upper limb functional use in breast cancer survivors [56,57]. Training strategies and hyperparameters followed the original studies of Pavllo et al. [31], and a testing procedure is outlined in Section Experimental setting. The implementation was carried out using Python version 3.12.2 and PyTorch version 2.2.1 and it is available at https://github.com/vayalet/Exercise-performance-DL.

Problem statement. The 2D pose sequence used as input to our framework can be represented as $\mathbf{X} \in \mathbb{R}^{T \times J \times C_{\text{in}}}$, where T denotes the number of frames in the sequence and J is the number of joints ($J = 17$). The input channels are represented by $C_{\text{in}} = 3$, corresponding to the *x* and *y* coordinates and confidence score for each joint. Each sequence $\mathbf{X}$ is divided into overlapping windows of fixed length k, and it is associated with a ground truth label vector $\mathbf{Y}$, which varies depending on the task. To generate predictions for each window i, the model learns a function $f : \mathbf{X}_i \rightarrow \hat{\mathbf{Y}}_i$ that maps the

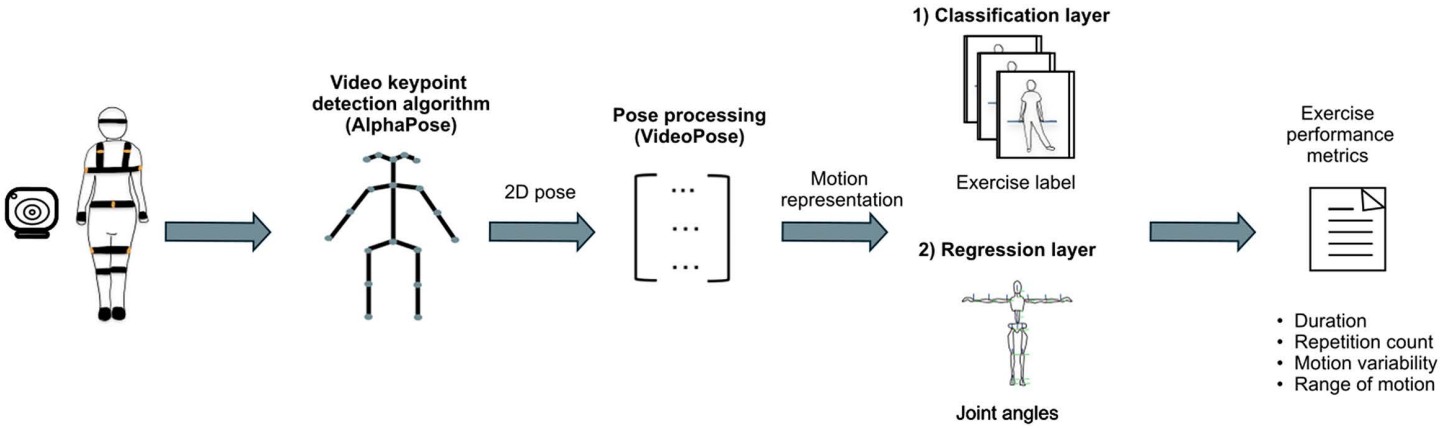

Fig 2. Overview of the workflow proposed in this study. Data is captured using a single camera and Xsens motion sensors. The AlphaPose algorithm extracts 2D poses from the video data, which are then processed by the VideoPose model to generate a motion representation. This representation is subsequently passed to a task-specific layer for either exercise recognition or joint angle prediction. A set of exercise metrics are computed based on the outputs of the two networks.

https://doi.org/10.1371/journal.pdig.0001576.g002

input window $\mathbf{X}_i = [\mathbf{x}_i^0, \ldots, \mathbf{x}_i^{k-1}]$ to an output $\hat{\mathbf{Y}}_i$ that is an approximation of the ground truth $\mathbf{Y}_i$ for that window. Each vector $\mathbf{x}_i^t \in \mathbb{R}^{J \times Cin}$ represents the 2D pose for frame t in the window. The task-specific details are outlined below:

- **Task 1:** The ground truth is a single action label corresponding to the central frame of each window, with $L = 15$ discrete action classes from the manual annotation of exercises. The model outputs a single label $\hat{\mathbf{Y}}_i \in \mathbb{R}^{1 \times L}$ per window of input data $\mathbf{X}_i$.
- **Task 2:** A sequence-to-sequence approach is implemented, where the ground truth corresponds to the 17 joint angles recorded with the motion sensors. Each keypoint is described by three angle types in terms of rotational degrees of freedom: x (lateral bending), y (flexion/extension), and z (axial rotation), which result in $\mathbf{Y}_i \in \mathbb{R}^{k \times J \times 3}$.

Deep learning network. The DL network implemented in this study was based on the VideoPose framework, which is a fully convolutional architecture proposed by Pavllo et al. [31]. This state-of-the-art model used temporal convolutions with ResNet-style blocks to capture temporal dependencies in 2D pose sequences. We followed the originally proposed pipeline, which outputs a 1×15 label prediction for Task 1 and a $1 \times 17 \times 3$ angle prediction for Task 2, both generated per window of 2D pose data. All hyperparameters were retained from the original architecture, including convolutional layers with 1024 output channels and a dropout rate of 0.25 [31]. Training was performed for 60 epochs with a batch size of 32. Following the protocol proposed by [31], the model was optimized using the Amsgrad variant of the Adam optimizer [58], with an initial learning rate of 10^{-3} that decayed exponentially by a factor of $\alpha = 0.95$ after each epoch. For the input data, the 2D pose data was split into overlapping windows of 243 frames with 50% overlap, aligning with the data characteristics that the VideoPose architecture was optimized for [31].

Task 1: To perform exercise recognition, the original regression model was adapted for classification by modifying the final convolutional layer to output exercise labels instead of continuous values and adjusting the output tensor shape to match classification targets. Cross-entropy loss was used to optimize the classification performance. To account for the class imbalance observed in our dataset, class weights were computed inversely proportional to class frequencies in the training set and incorporated into the loss function.

Task 2: For the regression task of estimating joint angles, we used a loss function that adapts the Mean Per-Joint Position Error (MPJPE) to angular space (MPJAE) by computing mean absolute differences between predicted and ground-truth joint angles across all joints and angular dimensions. The MPJAE loss is defined in Eq (1)

$$\Delta(\mathbf{y}, \mathbf{y}') = \frac{1}{J \cdot D} \sum_{j=1}^{J} \sum_{d=1}^{D} \left| \mathbf{y}_{j,d} - \mathbf{y}'_{j,d} \right|, \tag{1}$$

where $\mathbf{y}_{j,d}$ denotes the d-th rotational angle type ($D=3$) of joint j in pose $\mathbf{y}$, and J is the number of joints ($J=17$). Angular differences were normalized to the range $[0°, 180°]$ to account for circularity. Due to the reliability issues observed during sensor calibration and data collection, particularly noisy upper body joint kinematics, a subset of 9 joints was selected for this task. With this we ensured that only reliably recorded joint angle data was used as ground truth. A corresponding mask was applied to the loss function, focusing the training on the 3 axial joints (lumbosacral (L5S1), thoracolumbar (L1T12), and cervicothoracic (T1C7) [59]) and the lower body joints (left and right hip, knee, and ankle). This approach ensured that only the weights for reliable joints were updated, allowing the model to effectively learn the kinematics of clinically relevant angles.

Evaluation

To evaluate the performance of the models on the two tasks, a leave-one-subject-out (LOSO) cross-validation (CV) approach was implemented. In this method, the data is iteratively partitioned based on the number of subjects in the dataset. For each iteration, one subject is preserved for evaluation, while the data from the remaining subjects is used to train the model. This process is repeated until all subjects have been used for evaluation and the final results described for each of the two tasks represent the averaged performance across all iterations. The LOSO approach is particularly suitable for scenarios where models need to generalize well to new, unseen subjects, reflecting real-world applications [60] such as exercise assessment. However, the results are based on a small sample of seven participants, making this study more suitable as a proof-of-concept rather than for drawing broad conclusions.

Experimental setting. To compute the EPMs, predictions at a frame-by-frame resolution were required. During training, overlapping input sequences were used to maintain temporal continuity and use contextual information. However, for model inference, predictions were made for every frame in the input sequence ensuring a fine-grained resolution.

In Task 1 the classification model generated predictions for the central frame of each input window of training data. Therefore, during inference, test sequences were processed using a sliding window with a step size of 1 to ensure that every frame received a prediction. For consistency with the training procedure, the window size was set to 243 frames. An additional post-processing step was incorporated after the entire sequence for a subject was predicted. This step aimed to smooth out frame-level predictions by incorporating domain knowledge that each exercise typically lasts for at least 1 second, corresponding to 30 frames. Using this knowledge, if short sequences of frame-level predictions differed from the majority class within a window of size 30 frames, those frames were reassigned to the majority class of the window. This smoothing process reduced noise in the predictions and ensured more consistent and realistic classification outputs, aligning with the expected temporal dynamics of the exercises.

For the regression in Task 2, the architecture of VideoPose was designed to process the entire test sequence simultaneously, requiring no additional adjustments for inference [31]. The reported results were computed only for the pre-selected set of 3 axial and 6 lower body joints.

Metrics. To evaluate the performance of the described models on the two tasks, we used standard, task-appropriate evaluation metrics. These metrics are widely used in the literature across diverse populations, supporting their suitability for evaluating movement in older adults as well.

1) **Exercise Recognition:** Accuracy and segment-wise F1-score at k (F1@k) were used to evaluate model performance on the classification task. Due to its frame-by-frame nature, exercise recognition resembles an action segmentation task which, as proposed by Lea et al. [61], is better evaluated using segment-wise F1 scores rather than sample-wise F1 score. This is because it penalizes over- and under-segmentation errors by measuring how well predicted segments match ground truth segments with at least k% temporal overlap. For each predicted segment, which is a sequence of identical labels, the Intersection over Union (IoU) with ground-truth segments of the same class is calculated as in Eq (2):

$$\text{IoU} = \frac{\left|\text{Prediction} \cap \text{Ground Truth}\right|}{\left|\text{Prediction} \cup \text{Ground Truth}\right|}. \tag{2}$$

A predicted segment is considered a true positive (TP) if it overlaps with a ground-truth segment of the same class with an IoU $\geq k$. Predicted segments without a valid match are counted as false positives (FP), while unmatched ground-truth segments are false negatives (FN). Using these, Segment-F1@k is computed as in Eq (3):

$$\text{F1@}k = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN}}. \tag{3}$$

Following standard practice [61], we report results at Segment-F1@50 using a threshold of $k = 0.5$.

2) **Joint Angle Prediction:** The Mean Per-Joint Angular Error (MPJAE) was used to evaluate the regression task. This metric, defined in Eq (1), calculates the mean angular difference in degrees between the predicted and ground truth angles for all relevant joints. It provides a robust measure of the model's accuracy in predicting joint angles by capturing the average deviation across all frames and joints in the test sequence. Additionally, we report the Pearson correlation coefficient between the estimated and true angles as this is a common regression metric.

Exercise performance metrics (EPMs). To evaluate the performance of each physical exercise performed by the subjects, we calculated a set of EPMs specifically selected for their relevance for assessing physical function in older adults. The EPMs were computed using both model predictions and ground truth labels, and their reliability was assessed by calculating the Pearson correlation coefficient between the two. We chose to compute the EPMs of the fine-grained sub-exercise actions rather than the complete exercises (Table 1), as this level of detail can offer important insights into an older adult's performance. For example, taking longer to perform a lateral movement with the left leg compared to the right may suggest difficulties specific to one of the legs. Such distinctions can help identify which parts of an exercise are more physically challenging and may benefit from targeted intervention.

Total exercise duration: Monitoring exercise duration in older adults supports the evaluation of engagement in physical activity and can reveal changes in endurance associated with aging. The duration of each exercise was computed in terms of the cumulative time each subject spent performing a specific sub-exercise action. This was calculated by counting the total number of frames corresponding to each sub-exercise action label and converting it to minutes based on the video frame rate. Let $N_{\text{frames, }e}$ represent the number of frames labeled as sub-exercise action e. The total duration in minutes of sub-exercise action e, denoted as D_e, is calculated with Eq (4):

$$D_e = \frac{N_{\text{frames, }e}}{F \cdot 60}, \tag{4}$$

where F is the video frame rate (30 frames per second).

Number of repetitions: Repetition count reflects the frequency of movement and can help identify reduced activity in older adults linked to functional decline. The number of repetitions for each sub-exercise action was calculated by

identifying transitions in the frame-by-frame action labels. A new repetition was counted whenever the label transitioned from one sub-exercise action to another. To ensure robust counting, domain knowledge was incorporated to smooth out the label transition events. Specifically, a transition was only counted as a new repetition if the current label was the majority label in a window of 30 subsequent frames, corresponding to 1 second of activity. Formally, let L_t represent the label at frame t. A transition from frame t to frame $t+1$ contributes to the repetition count R_e for exercise e if:

$$\text{Majority}\left(\{L_t, L_{t+1}, \ldots, L_{t+29}\}\right) = e, \tag{5}$$

where Majority denotes the most frequent label within the 30-frame window. This smoothing process reduces noise in the label transitions and ensures only valid repetitions are counted.

Motion variability: Motion variability was included in this study as an important indicator of movement quality, neuromuscular control, and fall risk. Prior research has shown that increased variability in joint kinematics may reflect underlying stability impairments, particularly among older adults [46]. For example, high variability in hip flexion during tasks such as sitting, standing, or bending has been associated with reduced postural control and increased fall risk [62].

To quantify motion variability, we calculated the mean absolute deviation (MAD) of the rotational joint angles along the 3 degrees of freedom (i.e., lateral bending (x), flexion/extension (y), and axial bending (z)) during each exercise repetition. For a given joint j, angle $a \in \{x, y, z\}$, and exercise repetition r consisting of n time steps, MAD was defined in Eq (6):

$$\text{MAD}_{j,a}^{(r)} = \frac{1}{n}\sum_{t=1}^{n}\left|\theta_{j,a}{}^{(t)} - \bar{\theta}_{j,a}\right|, \tag{6}$$

where $\theta_{j,a}^{(t)}$ is the joint angle at time step t for joint j along rotational axis a, and $\bar{\theta}_{j,a}$ is the mean angle over the repetition. The time steps belonging to each exercise repetition were identified using the ground truth sub-exercise action labels from Task 1.

Range of motion (ROM): ROM was included as one of our EPMs due to its well-established relationship with age-related mobility decline and reduced physical function in older adults [63]. In particular, this loss in ROM has been shown to be joint-specific, likely reflecting the varying patterns of joint usage accumulated over a lifetime. Such joint-level insights can inform the personalization of exercise routines for older adults. In our study, ROM was defined as the difference between the maximum and minimum joint angle values observed within each exercise repetition. For a given joint j, rotational angle $a \in \{x, y, z\}$, and exercise repetition r consisting of n time steps, the ROM was computed with Eq (7):

$$\text{ROM}_{j,a}^{(r)} = \max_{1\le t\le n}\theta_{j,a}{}^{(t)} - \min_{1\le t\le n}\theta_{j,a}{}^{(t)}, \tag{7}$$

where $\theta_{j,a}{}^{(t)}$ represents the joint angle at time step t for joint j along rotational axis a.

Results

This section begins with a description of the data characteristics, followed by an overview of the performance of the implemented DL model on (1) exercise recognition and (2) joint angle prediction. Finally, the accuracy of EPMs estimated from model predictions is compared to those derived from ground truth data.

Dataset

The study included seven older adult participants, consisting of four males and three females, aged between 65 and 94 years (mean age: 82.4 ± 8.1 years). A summary of their demographic characteristics is presented in Table 2.

Table 2. Demographic characteristics.

Subject ID	Age range (years)	Height range (cm)	Gender	Weight (kg)
1	80–84	165–169	Female	111.8
2	65–69	170–174	Male	79.8
3	85–89	165–169	Female	52.2
4	85–89	170–174	Male	52.0
5	80–84	175–179	Male	91.6
6	90–94	160–164	Female	66.0
7	75–79	175–179	Male	70.0

Overview of height, age range, gender, and body weight at the time of data collection for each subject.

https://doi.org/10.1371/journal.pdig.0001576.t002

Table 3 provides an overview of the collected exercise data that was used in the analysis of this study. Each participant completed between 5 and 8 measurement sessions, resulting in a total duration ranging from 67 to 94 minutes per subject. The mean duration per session varied between 10 and 15 minutes which reflects differences in exercises pace and individual performance across subjects.

In addition, Table 4 summarizes the dataset across all subjects and provides an overview of the collected data for each sub-exercise action. When excluding the random movement (label 0), the exercises span from approximately 19–43 minutes. The shortest action corresponds to moving the right foot backward after touching a cup (label 12), while the longest one relates to leaning backward while seated on a chair (label 4). Similar variability is observed in the total number of repetitions, ranging from 1327 (label 11) to 1903 (label 4). This indicates an imbalance in the occurrence of the different exercises in the dataset, which is accounted for during training through the use of class weights in the loss function, as described in Section Deep learning network.

DL Model performance

We evaluated the performance of the DL architecture on both tasks: exercise recognition (Task 1) and joint angle estimation (Task 2). The results are summarized in Table 5 in terms of the mean LOSO CV performance metrics. To better understand the strengths and limitations of the model, we provide a detailed analysis of the performance in each task in the upcoming sections.

Exercise recognition (Task 1). From Fig 3 it can be seen that the model performed quite consistently in recognizing the different sub-exercise actions. The observed outliers for the last exercise of touching a cup with each foot, comprising

Table 3. Per subject dataset characteristics.

Subject	Measurements	Total Duration (min)	Mean Duration (min)
S1	5	67.33	13.47
S2	6	94.14	15.69
S3	6	90.72	15.12
S4	6	94.63	15.77
S5	5	71.98	14.40
S6	7	87.61	12.52
S7	8	82.53	10.32

Overview of the data collected for each subject including the number of measurement trials, the total duration of all measurements in minutes and the mean duration of a measurement in minutes.

https://doi.org/10.1371/journal.pdig.0001576.t003

Table 4. Per exercise dataset characteristics.

Exercise label	Mean total No. of frames	Total Duration (min)	No. of Repetitions
0	25747.4	100.13	752
1	10994.6	42.76	1377
2	12277.7	47.75	1372
3	10347.6	40.24	1900
4	11175.9	43.46	1903
5	8578.9	33.36	1475
6	8774.1	34.12	1491
7	10536.4	40.98	1565
8	10492.6	40.80	1576
9	9849.7	38.30	1782
10	9549.4	37.14	1775
11	5435.0	21.14	1327
12	5090.1	19.80	1328
13	5636.1	21.92	1352
14	5471.6	21.28	1349

Overview of the dataset at the sub-exercise action level, including the mean total number of frames, the total duration in minutes, and the total number of repetitions across all subjects. The exercise label corresponds to the sub-exercise action labels used in the analysis (see detailed description in Table 1).

https://doi.org/10.1371/journal.pdig.0001576.t004

Table 5. LOSO CV performance of the DL model on Task 1 (Exercise recognition) and Task 2 (joint angle estimation).

Task	Metric	LOSO CV mean (± SD)
Exercise recognition	F1@50	0.92 ± 0.04
	Accuracy	0.88 ± 0.04
Joint angle estimation	MPJAE [°]	7.7 ± 0.91
	Correlation (r)	0.75 ± 0.05

https://doi.org/10.1371/journal.pdig.0001576.t005

the last four sub-exercise actions, are due to one subject (Subject 2). During data collection this participant experienced particular difficulties with the exercises, exhibited restricted mobility, and required assistance during some of the movements. Additionally, the model struggled to classify random movements, likely due to their inherently higher motion variability.

Furthermore, the model output example of annotated time frames in Fig 4 illustrates that the model performed well in detecting the action boundaries in frame-by-frame annotation when compared to the manually annotated ground truth. However, it also highlights the challenge of accurately identifying the random movements that occur between exercises, likely because this label describes a wide variety of movement types. In this particular case, the subject may have been adjusting their position on the chair before initiating the next exercise, which the model interpreted as the bending backward action. Despite this, the overall classification of the complete exercise remains correct, as the model successfully identifies the transition to the bending forward/backward exercise sequence.

Joint angle estimation (Task 2). Fig 5 presents the LOSO CV results for estimating the three rotational angles across the 9 selected joints (3 axial and 6 lower body). Overall, the model demonstrates strong performance, though errors vary by joint. Notably, the hip and knee joints exhibit the highest estimation errors. This aligns with previous findings that highlighted greater variability and complexity in predicting lower body joint angles due to increased mobility and degrees of freedom [64].

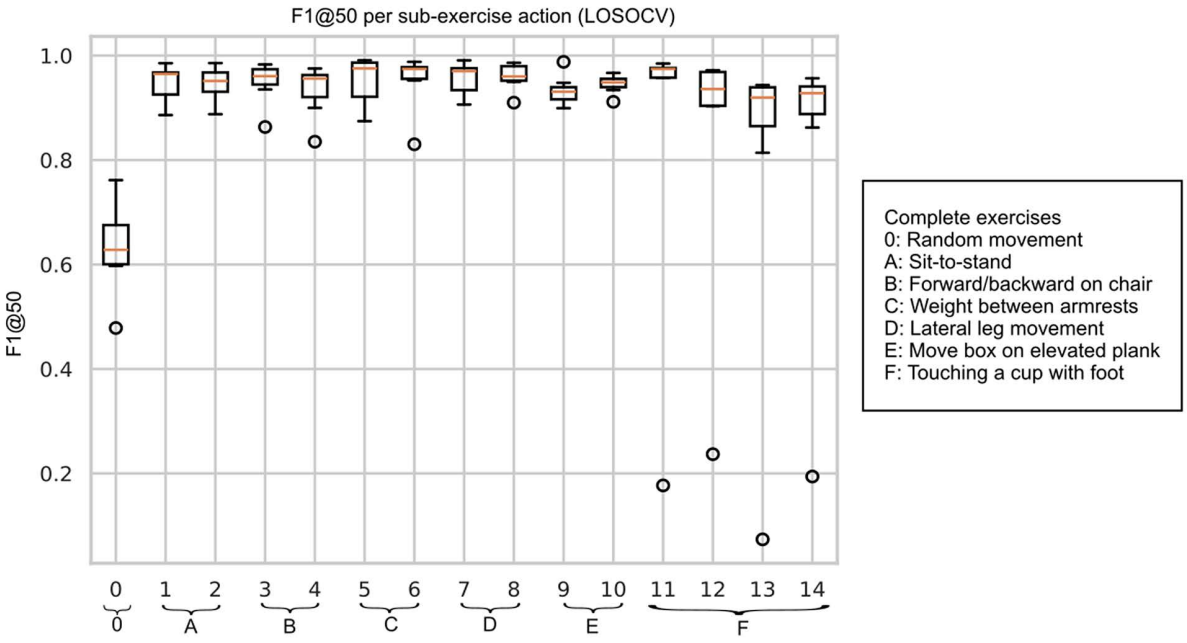

Fig 3. Model performance on exercise recognition (LOSO CV). Each boxplot represents the F1@50 scores per sub-exercise action for all subjects. For clarity the sub-exercise actions are grouped into their corresponding complete exercises (see Table 1 for detailed description on the annotations).

https://doi.org/10.1371/journal.pdig.0001576.g003

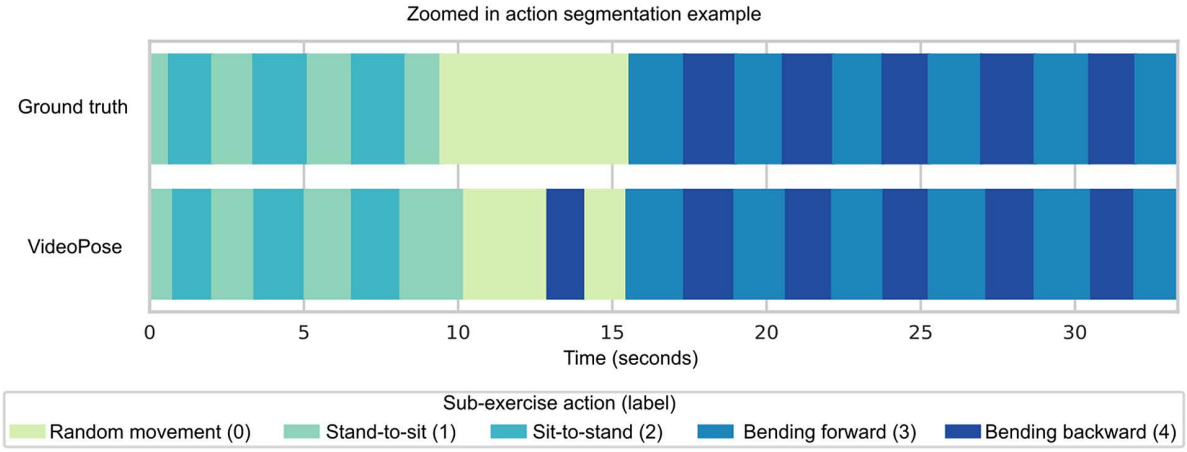

Fig 4. An example time segment of a subject trial with action segmentation. The figure displays sub-exercise action labels over time, comparing the manually annotated ground truth with predictions from our DL model. This example is from a transition where the subject switched from performing the sit-to-stand exercise to the forward and backward bending movement on a chair.

https://doi.org/10.1371/journal.pdig.0001576.g004

It is important to interpret these results within the context of the specific exercises, as the estimation errors differ depending on which joint is involved and how it contributes to the movement. To illustrate this in terms of the continuous joint angle signals generated by the model, Fig 6 presents a frame-wise example of an exercise sequence. In the first segment, where the subject is moving a weight between the armrests while seated, the hip flexion angle is not accurately estimated. However, this inaccuracy is less concerning, as hip flexion is not functionally relevant for this particular

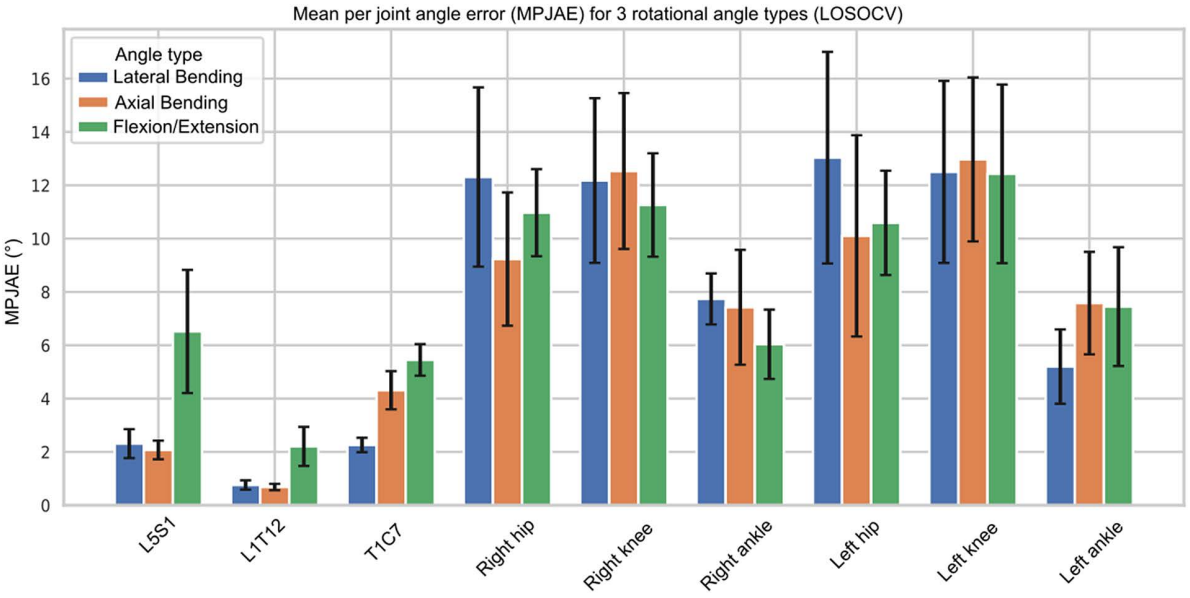

Fig 5. Model performance on joint angle prediction (LOSO CV). Each bar represents the mean MPJAE in degrees over all subjects and the error bars show the variability between LOSO CV folds. For each of the pre-selected 9 joints, the three rotational angles estimation errors are shown.

https://doi.org/10.1371/journal.pdig.0001576.g005

exercise. In contrast, the axial rotation of the L1T12 joint which is located along the thoracic-lumbar spine and is more engaged during this movement, is well estimated. Furthermore, in the subsequent sit-to-stand exercise, where hip flexion plays a central role in assessing movement quality and stability, the predicted signal aligns closely with the ground truth. This indicates strong performance of the model in estimating clinically relevant joint angles when interpreted in the context of the exercise.

Exercise performance metrics estimated from model predictions

The first two EPMs, total exercise duration and number of repetitions, were derived from the frame-level sub-exercise action labels in Task 1. As shown in Fig 7, the predicted exercise duration closely matches the ground truth, with a strong correlation coefficient of $r = 0.93$ with a 95% confidence interval (CI) [0.90, 0.95]. The deviations from the true values indicate that the model is more suitable for broader time estimations, such as on the scale of minutes, rather than highly precise measurements in the millisecond range. Similarly, the predicted repetition count strongly aligns with the true values, achieving a correlation of $r = 0.86$ with CI [0.80, 0.90]. Overall, the performance is consistent across exercises and subjects, with the exception of Subject 2, who had drops in action recognition performance, as previously detailed in Section Exercise recognition (Task 1).

The other two EPMs, motion variability (in terms of MAD) and ROM, were computed using the joint angle estimations from Task 2. As mentioned in the Task 2 results we expected to observe varying results in the calculation of these EPMs depending on the context in terms of joint, angle type and exercise. An example of the variability in results can be seen in Fig 8, where only the flexion angles are used to calculate MAD and ROM values. Strong correlations (up to 0.9) are observed for the hips and knees across multiple exercises. This is especially relevant for the sit-to-stand exercise (labels 1 and 2), which is among the most physically demanding ones and poses the highest stability risk for this population. In contrast, the ankles and the three axial joints (L5S1, L1T12, T1C7) consistently show lower performance in both MAD and ROM for the flexion/extension angle. This is likely because the corresponding exercises did not involve significant

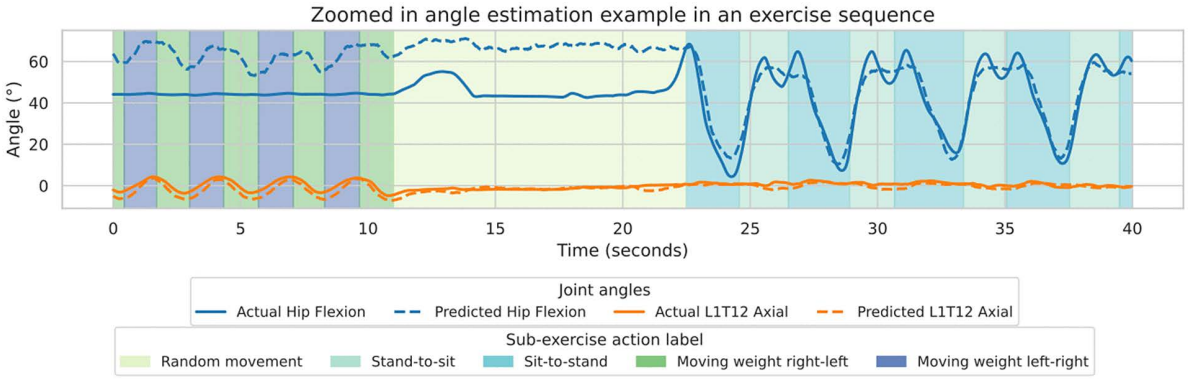

Fig 6. An example time segment of a subject trial with action segmentation and angle predictions. The example segment displays the predicted and ground truth hip flexion and L1T12 axial rotation angles during the transition between two exercises. In the background are the sub-exercise action annotations illustrated with different colors.

https://doi.org/10.1371/journal.pdig.0001576.g006

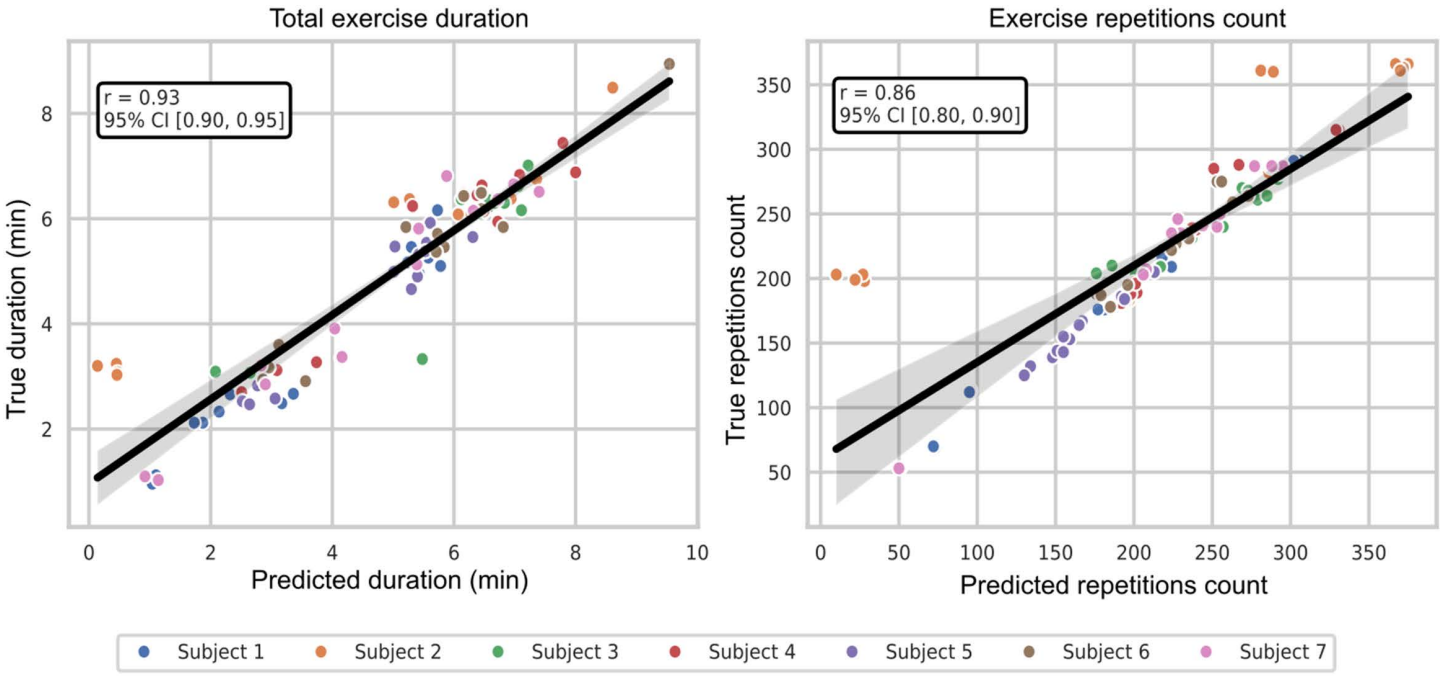

Fig 7. Overview of estimation performance of exercise duration and repetition count. Pearson's correlation coefficient (*r*) between predicted and actual exercise duration (left) and repetition count (right) across subjects. Each point represents a sub-exercise action, color-coded by subject. The solid black line indicates the linear regression fit, and the shaded area shows the 95% confidence interval.

https://doi.org/10.1371/journal.pdig.0001576.g007

flexion in those joints, making their motion patterns subtler and their estimations noisier, which led to noisier calculations of EPMs.

Our primary objective in computing these EPMs is to demonstrate their potential in supporting clinicians to monitor exercise performance in older adults over time. Fig 9 illustrates this by showing the evolution of MAD and ROM values for the hip and knee flexion angles across 8 repetitions of the sit-to-stand exercise for a single participant. In this example,

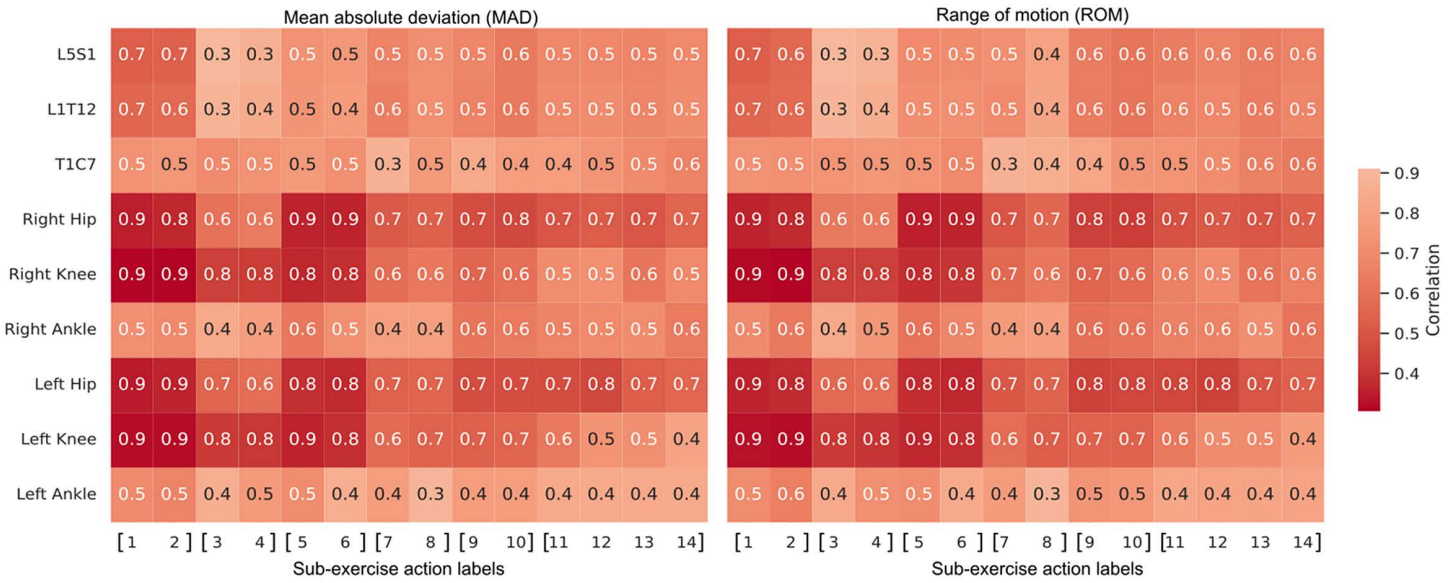

Fig 8. Overview of estimation performance of mean absolute deviation and range of motion. Correlation between the predicted and true mean absolute deviation (MAD, left) and range of motion (ROM, right) for the flexion/extension angle. Each cell represents the correlation for a specific joint (rows) and sub-exercise action (columns), averaged over all subjects. The action labels are grouped to their corresponding complete exercises using square brackets '[]'. The varying scores highlight the influence of joint function and exercise type on model performance.

https://doi.org/10.1371/journal.pdig.0001576.g008

the participant displays an overall stable hip and knee flexion movement, as reflected by relatively consistent MAD values. However, during two early repetitions (circled in red), both hip and knee ROM values are noticeably lower, indicating the subject did not fully extend their joints when standing up. The reduced movement is also accompanied by low MAD scores, which does not seem to be due to improved movement variability, but rather due to limited ROM. This highlights the importance of interpreting MAD alongside ROM to avoid misleading conclusions. For instance, a consistent decline in both ROM and MAD over time might suggest that the participant is no longer exercising using their full motor capabilities. This could be either due to emerging physical pain or due to a lack of motivation. In contrast a low ROM coupled with high MAD values might indicate stability or balance issues. Identifying such patterns can facilitate an intervention by the caregivers allowing them to adapt the exercise program to better suit the participant's evolving needs.

Discussion

This study presents a novel method for the unsupervised monitoring of older adults' exercise performance using a single-camera system combined with deep learning. We demonstrated that a DL model can be trained to successfully recognize exercise actions and estimate joint angles in older adults. Our implementation relied on existing tools: AlphaPose for 2D pose estimation and VideoPose for processing temporal pose data for both classification and regression tasks. While the system showed promising performance, future work could explore whether alternative 2D pose estimators, such as OpenPose [65] and PoseNet [66], or different model architectures might improve accuracy. However, it is important that any future improvements take into account the computational power required to run these algorithms to ensure that their deployment in real-world scenarios is practical.

A key contribution of this study is the use of 3D joint angles for exercise monitoring, rather than the commonly used 3D joint coordinates. By focusing on angles, our approach offers a more interpretable and task-relevant measure of joint movement, particularly for assessing outcomes like motion variability and range of motion, highlighting their value for

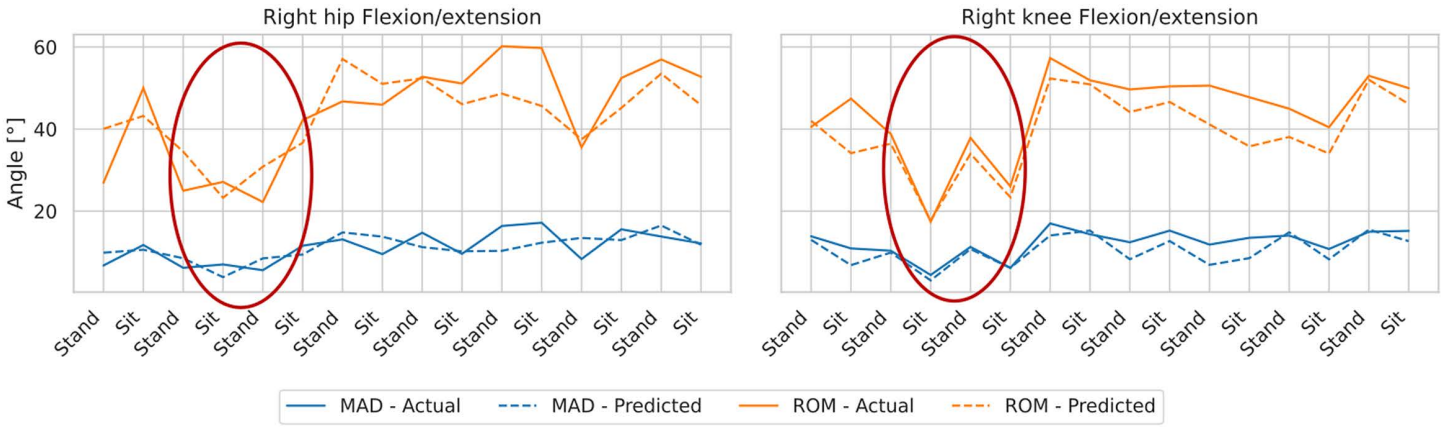

Fig 9. Example segment of the sit-to-stand exercise with MAD and ROM values. The segment displays the predicted and actual MAD (blue) and ROM (orange) values for the flexion/extension angle of the right hip (left) and ankle (right). The figure shows the changes in both EPMs over time during each sub-exercise action of 8 repetitions of the complete sit-to-stand exercise. The part highlighted with a red eclipse is a short segment where the participant did not stand up as high as in other repetitions, but quickly corrected themselves.

https://doi.org/10.1371/journal.pdig.0001576.g009

monitoring older adults' exercises. The upper body joints were excluded from the analysis due to difficulties in data collection, such as inconsistent calibration of the upper body skeleton, which introduced noise into the motion data. This limitation highlights the challenges of capturing precise motion data in older adults, particularly for exercises involving subtle or limited joint movements. Another potential concern could be our use of motion sensors as a ground truth for joint angles, which are known to have a degree of measurement inaccuracy [67]. However, this is not a major issue for our application, which focuses on tracking changes over time rather than clinical evaluation of individual exercise execution. Even with minor sensor inaccuracies, longitudinal trends in ROM and motion variability still provide meaningful insights into functional progression.

Additionally, this work focused on a few key exercise outcomes, but more metrics, such as limb coordination or muscle power, could be incorporated in future studies. Our system offers significant potential for longitudinal monitoring of exercise performance due to the contactless nature of the data. This could enable caretakers in residential homes to adjust exercise programs to better fit the individual needs of older adults. The approach aligns with the WHO's recommendations for healthy aging, which emphasize the importance of maintaining functional ability through regular physical activity and personalized exercise interventions [23]. The framework could also be adapted for real-time processing to provide corrective feedback during exercises, further improving exercise quality and safety [68]. Importantly, the method underscores the potential of our system to provide meaningful insights into exercise performance in accordance with established guidelines for healthy aging. This method offers a cost-effective solution for care homes to facilitate regular exercise sessions for residents without requiring active caregiver supervision. Staffing is one of the primary costs associated with care homes in Europe [69], and this approach could help minimizing labor costs by utilizing a simple webcam with no high-resolution requirements, reducing the need for expensive equipment or additional personnel.

A limitation of this study is the small sample size and the fact that data collection was conducted in a single facility, which may restrict the generalization of the findings. However, the sample captured the heterogeneity within this population, including varying levels of frailty and differences in movement patterns. The controlled environment of the facility, with minimal visual distractions such as paintings, plants, or other objects, may have facilitated the accuracy of the motion tracking system. Future studies should explore whether the approach remains effective in more visually complex environments, including home settings, where factors like furniture, lighting conditions, and background movement could impact tracking performance. Finally, our current approach is limited to single-person analysis. Extending the system to group

exercise settings could enhance its applicability by enabling interactive sessions for older adults, fostering engagement and social interaction during physical activity [70].

Conclusion

This study aimed to address the need for scalable and cost-effective solutions for unsupervised monitoring of older adults' exercise performance. By leveraging an existing model architecture for 3D pose estimation (VideoPose) we successfully adapted it to recognize exercises and estimate rotational angles for axial and lower body joints, enabling the assessment of key exercise performance metrics such as duration, repetition count, motion variability, and range of motion.

This work highlights the potential of a deep learning-based system to monitor exercise performance in older adults, offering opportunities for longitudinal tracking to support tailored physical activity programs. Future improvements, such as enabling real-time processing and extending the system to group exercise analysis, could further enhance its utility in promoting the health and independence of older adults.

Acknowledgments

We thank all the participants who generously gave their time and effort to contribute to this study.

Author contributions

Conceptualization: Vayalet Stefanova, Dimitri Vargemidis, Julien Lebleu, David Beckwée, Moran Gilat, Benjamin Filtjens, Bart Vanrumste.

Data curation: Evelien Maeyens, Jeroen Brughmans, Karen Gilis, Benjamin Filtjens.

Funding acquisition: Benjamin Filtjens, Bart Vanrumste.

Investigation: Vayalet Stefanova.

Methodology: Vayalet Stefanova, David Beckwée, Benjamin Filtjens, Bart Vanrumste.

Project administration: Leen Stulens, Benjamin Filtjens, Bart Vanrumste.

Resources: Bart Vanrumste.

Software: Vayalet Stefanova.

Supervision: Moran Gilat, Benjamin Filtjens, Bart Vanrumste.

Validation: Vayalet Stefanova.

Visualization: Vayalet Stefanova.

Writing – original draft: Vayalet Stefanova, Benjamin Filtjens.

Writing – review & editing: Vayalet Stefanova, Evelien Maeyens, Jeroen Brughmans, Julien Lebleu, Leen Stulens, Sanne Broeder, Karen Gilis, Mieke Deschodt, Benedicte Vanwanseele, David Beckwée, Moran Gilat, Benjamin Filtjens, Bart Vanrumste.

References

1. Population Division D. World Population Prospects 2024: Summary of Results — DESA Publications [Internet]. desapublications.un.org. United Nations: Department of Economic and Social Affairs; 2024. Available from: https://desapublications.un.org/publications/world-population-prospects-2024-summary-results
2. McMurdo ME, Rennie L. A controlled trial of exercise by residents of old people's homes. Age Ageing. 1993;22(1):11–5. https://doi.org/10.1093/ageing/22.1.11 PMID: 8438659
3. Salguero A, Martínez-García R, Molinero O, Márquez S. Physical activity, quality of life and symptoms of depression in community-dwelling and institutionalized older adults. Arch Gerontol Geriatr. 2011;53(2):152–7. https://doi.org/10.1016/j.archger.2010.10.005 PMID: 21035206

4. Bean JF, Vora A, Frontera WR. Benefits of exercise for community-dwelling older adults. Arch Phys Med Rehabil. 2004;85(7 Suppl 3):S31-42; quiz S43-4. https://doi.org/10.1016/j.apmr.2004.03.010 PMID: 15221722
5. Domingos C, Pêgo JM, Santos NC. Effects of physical activity on brain function and structure in older adults: a systematic review. Behav Brain Res. 2021;402:113061. https://doi.org/10.1016/j.bbr.2020.113061 PMID: 33359570
6. Izquierdo M, Merchant RA, Morley JE, Anker SD, Aprahamian I, Arai H, et al. International exercise recommendations in older adults (ICFSR): expert consensus guidelines. J Nutr Health Aging. 2021;25(7):824–53. https://doi.org/10.1007/s12603-021-1665-8 PMID: 34409961
7. Baldelli G, De Santi M, De Felice F, Brandi G. Physical activity interventions to improve the quality of life of older adults living in residential care facilities: a systematic review. Geriatr Nurs. 2021;42(4):806–15. https://doi.org/10.1016/j.gerinurse.2021.04.011 PMID: 34090224
8. Phillips EM, Schneider JC, Mercer GR. Motivating elders to initiate and maintain exercise. Arch Phys Med Rehabil. 2004;85(7 Suppl 3):S52-7; quiz S58-9. https://doi.org/10.1016/j.apmr.2004.03.012 PMID: 15221728
9. Peek K, Sanson-Fisher R, Mackenzie L, Carey M. Interventions to aid patient adherence to physiotherapist prescribed self-management strategies: a systematic review. Physiotherapy. 2016;102(2):127–35. https://doi.org/10.1016/j.physio.2015.10.003 PMID: 26821954
10. de Souto Barreto P, Morley JE, Chodzko-Zajko W, H Pitkala K, Weening-Djiksterhuis E, Rodriguez-Mañas L, et al. Recommendations on physical activity and exercise for older adults living in long-term care facilities: a taskforce report. J Am Med Dir Assoc. 2016;17(5):381–92. https://doi.org/10.1016/j.jamda.2016.01.021 PMID: 27012368
11. Clarke N, Smith R, Wood J, Koskela S, Jones F, Hurley M. A qualitative interview study comparing and contrasting resident and staff perspectives of engaging in meaningful activity in a UK care home. Arch Gerontol Geriatr. 2019;83:257–62. https://doi.org/10.1016/j.archger.2019.05.005 PMID: 31103919
12. Brox E, Konstantinidis ST, Evertsen G, Fernandez-Luque L, Remartinez A, Oesch P, et al. Gameup: Exergames for mobility–a project to keep elderly active. Springer; 2016. p. 1225–30.
13. Swinnen N, Vandenbulcke M, de Bruin ED, Akkerman R, Stubbs B, Firth J, et al. The efficacy of exergaming in people with major neurocognitive disorder residing in long-term care facilities: a pilot randomized controlled trial. Alzheimers Res Ther. 2021;13(1):70. https://doi.org/10.1186/s13195-021-00806-7 PMID: 33785077
14. Crespo AB, Idrovo GG, Rodrigues N, Pereira A. A virtual reality UAV simulation with body area networks to promote the elders life quality. IEEE; 2016. p. 1–7.
15. Kimura K, Hozumi N. Investigating the acute effect of an aerobic dance exercise program on neuro-cognitive function in the elderly. Psychol Sport Exerc. 2012;13(5):623–9.
16. Douka S, Zilidou VI, Lilou O, Manou V. Traditional dance improves the physical fitness and well-being of the elderly. Front Aging Neurosci. 2019;11:75. https://doi.org/10.3389/fnagi.2019.00075 PMID: 31024290
17. Görer B, Salah AA, Akın HL. A robotic fitness coach for the elderly. In: Ambient Intelligence: 4th International Joint Conference, AmI 2013; 2013 Dec 3–5; Dublin, Ireland, Proceedings 4. Springer; 2013. p. 124–39.
18. Avioz-Sarig O, Olatunji S, Sarne-Fleischmann V, Edan Y. Robotic system for physical training of older adults. Int J Soc Robot. 2021;13(5):1109–24. https://doi.org/10.1007/s12369-020-00697-y PMID: 33020706
19. Fasola J, Matarić MJ. A socially assistive robot exercise coach for the elderly. J Hum-Robot Interact. 2013;2(2):3–32.
20. Lee MH, Siewiorek DP, Smailagic A, Bernardino A, Badia SBI. Enabling AI and robotic coaches for physical rehabilitation therapy: iterative design and evaluation with therapists and post-stroke survivors. Int J Soc Robot. 2024;16(1):1–22.
21. Bavan L, Surmacz K, Beard D, Mellon S, Rees J. Adherence monitoring of rehabilitation exercise with inertial sensors: a clinical validation study. Gait Posture. 2019;70:211–7. https://doi.org/10.1016/j.gaitpost.2019.03.008 PMID: 30903993
22. Broeder S, Filtjens B, Stulens L, Vargemidis D, Stefanova V, Vanwanseele B, et al. Perspectives of older adults and their caregivers on the development of a robotic exercise coach. Gerontechnology. 2024;23(2):1–1.
23. Organization WH. World report on ageing and health [Internet]. Geneva, Switzerland: World Health Organization; 2015. Available from: https://iris.who.int/bitstream/handle/10665/186463/9789240694811_eng.pdf?sequence=1
24. Filtjens B, Amsters R, Bin Junaid A, Damen N, Van De Laer J, Vanwanseele B, et al. Vision-based marker-less spatiotemporal gait analysis by using a mobile platform: preliminary validation. Springer; 2019. p. 126–41.
25. Xu X, McGorry RW, Chou L-S, Lin J-H, Chang C-C. Accuracy of the Microsoft Kinect for measuring gait parameters during treadmill walking. Gait Posture. 2015;42(2):145–51. https://doi.org/10.1016/j.gaitpost.2015.05.002 PMID: 26002604
26. Uhlrich SD, Falisse A, Kidziński Ł, Muccini J, Ko M, Chaudhari AS, et al. OpenCap: Human movement dynamics from smartphone videos. PLoS Comput Biol. 2023;19(10):e1011462. https://doi.org/10.1371/journal.pcbi.1011462 PMID: 37856442
27. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–44. https://doi.org/10.1038/nature14539 PMID: 26017442
28. Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, et al. Microsoft coco: common objects in context. Computer Vision–ECCV 2014: 13th European Conference; 2014 Sept 6–12; Zurich, Switzerland, Proceedings, Part V 13. Springer; 2014. p. 740–55.
29. Cao Z, Hidalgo G, Simon T, Wei S-E, Sheikh Y. OpenPose: realtime multi-person 2D pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell. 2021;43(1):172–86. https://doi.org/10.1109/TPAMI.2019.2929257 PMID: 31331883

30. Fang H-S, Li J, Tang H, Xu C, Zhu H, Xiu Y, et al. AlphaPose: whole-body regional multi-person pose estimation and tracking in real-time. IEEE Trans Pattern Anal Mach Intell. 2023;45(6):7157–73. https://doi.org/10.1109/TPAMI.2022.3222784 PMID: 37145952
31. Pavllo D, Feichtenhofer C, Grangier D, Auli M. 3D human pose estimation in video with temporal convolutions and semi-supervised training. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2019. p. 7745–54.
32. Mehraban S, Adeli V, Taati B. MotionAGFormer: enhancing 3D human pose estimation with a Transformer-GCNFormer network. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision; 2024. p. 6920–30.
33. Zhu W, Ma X, Liu Z, Liu L, Wu W, Wang Y. MotionBERT: a unified perspective on learning human motion representations. Proceedings of the IEEE/CVF International Conference on Computer Vision; 2023. p. 15085–99.
34. Filtjens B, Vanrumste B, Slaets P. Skeleton-based action segmentation with multi-stage spatial-temporal graph convolutional neural networks. IEEE Trans Emerg Top Comput. 2024;12(1):202–12.
35. Lea C, Vidal R, Hager GD. Learning convolutional action primitives for fine-grained action recognition. 2016 IEEE International Conference on Robotics and Automation (ICRA). IEEE; 2016. p. 1642–9.
36. Buzzelli M, Albé A, Ciocca G. A vision-based system for monitoring elderly people at home. Appl Sci. 2020;10(1):374. https://doi.org/10.3390/app10010374
37. Seethapathi N, Wang S, Saluja R, Blohm G, Kording KP. Movement science needs different pose tracking algorithms. arXiv:1907.10226 [Preprint]. 2019 [cited 2019 Jul 24]. Available from: https://doi.org/10.48550/arXiv.1907.10226
38. Sardari S, Sharifzadeh S, Daneshkhah A, Nakisa B, Loke SW, Palade V, et al. Artificial Intelligence for skeleton-based physical rehabilitation action evaluation: a systematic review. Comput Biol Med. 2023;158:106835. https://doi.org/10.1016/j.compbiomed.2023.106835 PMID: 37019012
39. Shukla BK, Jain H, Vijay V, Yadav SK, Mathur A, Hewson DJ. A comparison of four approaches to evaluate the sit-to-stand movement. IEEE Trans Neural Syst Rehabil Eng. 2020;28(6):1317–24. https://doi.org/10.1109/TNSRE.2020.2987357 PMID: 32310776
40. Lee C-H, Mendoza T, Huang C-H, Sun T-L. Vision-based postural balance assessment of sit-to-stand transitions performed by younger and older adults. Gait Posture. 2025;117:245–53. https://doi.org/10.1016/j.gaitpost.2025.01.001 PMID: 39798419
41. Mehdizadeh S, Nabavi H, Sabo A, Arora T, Iaboni A, Taati B. Concurrent validity of human pose tracking in video for measuring gait parameters in older adults: a preliminary analysis with multiple trackers, viewing angles, and walking directions. J Neuroeng Rehabil. 2021;18(1):139. https://doi.org/10.1186/s12984-021-00933-0 PMID: 34526074
42. Ng K-D, Mehdizadeh S, Iaboni A, Mansfield A, Flint A, Taati B. Measuring gait variables using computer vision to assess mobility and fall risk in older adults with dementia. IEEE J Transl Eng Health Med. 2020;8:2100609. https://doi.org/10.1109/JTEHM.2020.2998326 PMID: 32537265
43. Romeo L, Marani R, D'Orazio T, Cicirelli G. Video based mobility monitoring of elderly people using deep learning models. IEEE Access. 2023;11:2804–19. https://doi.org/10.1109/access.2023.3234421
44. Yu X, Xiong S. A dynamic time warping based algorithm to evaluate kinect-enabled home-based physical rehabilitation exercises for older people. Sensors (Basel). 2019;19(13):2882. https://doi.org/10.3390/s19132882 PMID: 31261746
45. Power V, Clifford AM. Characteristics of optimum falls prevention exercise programmes for community-dwelling older adults using the FITT principle. Eur Rev Aging Phys Act. 2013;10(2):95–106. https://doi.org/10.1007/s11556-012-0108-2
46. Hamacher D, Singh NB, Van Dieën JH, Heller MO, Taylor WR. Kinematic measures for assessing gait stability in elderly individuals: a systematic review. J R Soc Interface. 2011;8(65):1682–98. https://doi.org/10.1098/rsif.2011.0416 PMID: 21880615
47. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. JAMA. 1963;185:914–9.
48. van Dijk M, Allegaert P, Locus M, Deschodt M, Verheyden G, Tournoy J, et al. Geriatric Activation Program Pellenberg, a novel physiotherapy program for hospitalized patients on a geriatric rehabilitation ward. Physiother Res Int. 2021;26(3):e1905. https://doi.org/10.1002/pri.1905 PMID: 33783923
49. Schepers M, Giuberti M, Bellusci G, et al. Xsens MVN: consistent tracking of human motion using inertial sensing. Xsens Technol. 2018;1(8):1–8.
50. Robert-Lachaine X, Mecheri H, Larue C, Plamondon A. Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis. Med Biol Eng Comput. 2017;55(4):609–19. https://doi.org/10.1007/s11517-016-1537-2 PMID: 27379397
51. Luvizon DC, Picard D, Tabia H. 2D/3D Pose estimation and action recognition using multitask deep learning. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2018.
52. Choi H, Moon G, Lee KM. Pose2Mesh: Graph convolutional network for 3D human pose and mesh recovery from a 2D human pose. Springer International Publishing; 2020. p. 769–87.
53. Chen B, Ji H, Wang Z, Filtjens B, Wang C, Ren W, et al. Language-assisted human part motion learning for skeleton-based Temporal Action Segmentation. arXiv:241006353 [Preprint]. 2024. Available from: https://arxiv.org/abs/2410.06353
54. Stenum J, Cherry-Allen KM, Pyles CO, Reetzke RD, Vignos MF, Roemmich RT. Applications of pose estimation in human health and performance across the lifespan. Sensors (Basel). 2021;21(21):7315. https://doi.org/10.3390/s21217315 PMID: 34770620
55. Jin S, Xu L, Xu J, Wang C, Liu W, Qian C, et al. Whole-body human pose estimation in the wild. Springer; 2020. p. 196–214.
56. Yang P-K, Filtjens B, Ginis P, Goris M, Nieuwboer A, Gilat M, et al. Automatic detection and assessment of freezing of gait manifestations. IEEE Trans Neural Syst Rehabil Eng. 2024;32:2699–708. https://doi.org/10.1109/TNSRE.2024.3431208 PMID: 39028610

57. Emmerzaal J, Filtjens B, Vets N, Vanrumste B, Smeets A, De Groef A, et al. A data-driven approach to detect upper limb functional use during daily life in breast cancer survivors using wrist-worn sensors. Sci Rep. 2024;14(1):18165. https://doi.org/10.1038/s41598-024-67497-6 PMID: 39107354

58. Reddi SJ, Kale S, Kumar S. On the Convergence of Adam and Beyond. arXiv:1904.09237 [Preprint]; 2019. Available from: https://arxiv.org/abs/1904.09237

59. Movella. MVN Biomechanical Model [Internet]; 2023. Available from: https://base.movella.com/s/article/MVN-Biomechanical-Model?language=en_US

60. Saeb S, Lonini L, Jayaraman A, Mohr DC, Kording KP. The need to approximate the use-case in clinical machine learning. Gigascience. 2017;6(5):1–9. https://doi.org/10.1093/gigascience/gix019 PMID: 28327985

61. Lea C, Flynn MD, Vidal R, Reiter A, Hager GD. Temporal convolutional networks for action segmentation and detection. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017. p. 1003–12. Available from: https://doi.ieeecomputersociety.org/10.1109/CVPR.2017.113

62. Granacher U, Gollhofer A, Hortobágyi T, Kressig RW, Muehlbauer T. The importance of trunk muscle strength for balance, functional performance, and fall prevention in seniors: a systematic review. Sports Med. 2013;43(7):627–41. https://doi.org/10.1007/s40279-013-0041-1 PMID: 23568373

63. Medeiros HBdO, Araújo DSMSd, Araújo CGSd. Age-related mobility loss is joint-specific: an analysis from 6,000 flexitest results. Age. 2013;35(6):2399–407. https://doi.org/10.1007/s11357-013-9525-z

64. Sarvestan J, Ataabadi PA, Yazdanbakhsh F, Abbasi S, Abbasi A, Svoboda Z. Lower limb joint angles and their variability during uphill walking. Gait Posture. 2021;90:434–40. https://doi.org/10.1016/j.gaitpost.2021.09.195 PMID: 34597985

65. Cao Z, Hidalgo G, Simon T, Wei S-E, Sheikh Y. OpenPose: realtime multi-person 2D pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell. 2021;43(1):172–86. https://doi.org/10.1109/TPAMI.2019.2929257 PMID: 31331883

66. Papandreou G, Zhu T, Chen LC, Gidaris S, Tompson J, Murphy K. Personlab: Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model. Proceedings of the European Conference on computer vision (ECCV); 2018. p. 269–86.

67. Poitras I, Dupuis F, Bielmann M, Campeau-Lecours A, Mercier C, Bouyer LJ, et al. Validity and reliability of wearable sensors for joint angle estimation: a systematic review. Sensors (Basel). 2019;19(7):1555. https://doi.org/10.3390/s19071555 PMID: 30935116

68. Yi J, Lee S, Lee SH. The effect of the Smart Health Continuous Feedback For Elderly Exercise (SHe CoFFEE) program on mobility: a randomized controlled pilot study. Front Public Health. 2024;12:1442064. https://doi.org/10.3389/fpubh.2024.1442064 PMID: 39346580

69. Molinuevo D, Anderson R. Care homes for older Europeans: public, for-profit and non-profit providers — European Foundation for the Improvement of Living and Working Conditions. European Foundation for the Improvement of Living and Working Conditions; 2017. https://doi.org/10.2806/737365

70. Liu Y, Lachman ME. A group-based walking study to enhance physical activity among older adults: the role of social engagement. Res Aging. 2021;43(9–10):368–77. https://doi.org/10.1177/0164027520963613 PMID: 33021146