FEDERAL UNIVERSITY OF UBERLÂNDIA
FACULTY OF ELECTRICAL ENGINEERING
POSTGRADUATE PROGRAM IN ELECTRICAL ENGINEERING

AI-DRIVEN LIGHTWEIGHT MOTOR ASSESSMENT EMBEDDED IN AN EXERGAME FOR POST-STROKE REHABILITATION

JÚLIA TANNÚS DE SOUZA

UBERLÂNDIA, BRAZIL
2026

JÚLIA TANNÚS DE SOUZA

AI-DRIVEN LIGHTWEIGHT MOTOR ASSESSMENT EMBEDDED IN AN EXERGAME FOR POST-STROKE REHABILITATION

Thesis presented to the Postgraduate Program in Electrical Engineering of the Faculty of Electrical Engineering at the Federal University of Uberlândia in partial fulfillment of the requirements for obtaining the degree of Doctor of Science.

Concentration area: Information Processing

Advisor: Prof. Dr. Luiz Carlos Gomes de Freitas

Co-advisor: Prof. Dr. Eduardo Lázaro Martins Naves

UBERLÂNDIA, BRAZIL

2026

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

S729
2026

Souza, Júlia Tannús de, 1995-
AI-driven lightweight motor assessment embedded in an exergame for post-stroke rehabilitation [recurso eletrônico] : . / Júlia Tannús de Souza. - 2026.

Orientador: Luiz Carlos Gomes de Freitas.
Coorientador: Eduardo Lázaro Martins Naves.
Tese (Doutorado) - Universidade Federal de Uberlândia, Pós-graduação em Engenharia Elétrica.
Modo de acesso: Internet.
DOI http://doi.org/10.14393/ufu.te.2026.229
Inclui bibliografia.
Inclui ilustrações.

1. Engenharia elétrica. I. Freitas, Luiz Carlos Gomes de,1976-, (Orient.). II. Naves, Eduardo Lázaro Martins,1970-, (Coorient.). III. Universidade Federal de Uberlândia. Pós-graduação em Engenharia Elétrica. IV. Título.

CDU: 621.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:
Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074

JÚLIA TANNÚS DE SOUZA

AI-DRIVEN LIGHTWEIGHT MOTOR ASSESSMENT EMBEDDED IN AN EXERGAME FOR POST-STROKE REHABILITATION

Thesis presented to the Postgraduate Program in Electrical Engineering of the Faculty of Electrical Engineering at the Federal University of Uberlândia in partial fulfillment of the requirements for obtaining the degree of Doctor of Science.

Concentration area: Information Processing

Examination Committee:

Prof. Dr. Éderson Rosa da Silva – Federal University of Uberlândia (Chair)

Committee Members:

1. Prof. Dr. Eduardo Lázaro Martins Naves – Federal University of Uberlândia (Co-advisor)
2. Prof. Guy Bourhis, PhD – Université de Lorraine, France
3. Prof. Yann Morère, PhD – Université de Lorraine, France
4. Prof. Pierre Pino, PhD – Université de Lorraine, France
5. Prof. Thiago Braga Rodrigues, PhD – Technological University of the Shannon, Ireland

UBERLÂNDIA, BRAZIL

2026

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Engenharia Elétrica
Av. João Naves de Ávila, 2121, Bloco 3N - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
Telefone: (34) 3239-4707 - www.posgrad.feelt.ufu.br - copel@ufu.br

ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Engenharia Elétrica				
Defesa de:	Tese de Doutorado, 357, PPGEELT				
Data:	Dois de março de dois mil e vinte e seis	Hora de início:	10:00	Hora de encerramento:	12:30
Matrícula do Discente:	12213EEL004				
Nome do Discente:	Júlia Tannús de Souza				
Título do Trabalho:	AI-driven lightweight motor assessment embedded in an exergame for post-stroke rehabilitation				
Área de concentração:	Processamento da Informação				
Linha de pesquisa:	Metodologia e Técnicas da Computação				
Projeto de Pesquisa de vinculação:	Coordenador do projeto: Eduardo Lázaro Martins Naves. Título do Projeto: GT-Reabnet: Rede de Telereabilitação por meio de Realidade Virtual e Realidade Aumentada. Agência financiadora: Rede Nacional de Ensino e Pesquisa (RNP). Número do processo na agência financiadora: não se aplica. Vigência do projeto: 01/2021 a 12/2024.				

Reuniu-se através de videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Elétrica, assim composta:

Doutores: Guy Bourhis (Université de Lorraine), Yann Morère (Université de Lorraine), Pierre Pino (Université de Lorraine), Thiago Braga Rodrigues (Technological University of the Shannon), Eduardo Lázaro Martins Naves, coorientador da discente e Ederson Rosa da Silva, presidente da mesa.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Ederson Rosa da Silva, apresentou a Comissão Examinadora e a candidata, agradeceu a presença do público, e concedeu à discente a palavra para a exposição do seu trabalho. A duração da apresentação da discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir, o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir a candidata. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando a candidata:

APROVADA.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor. O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme, foi assinada pela Banca Examinadora.

Documento assinado eletronicamente por **MORERE**, **Usuário Externo**, em 02/03/2026, às 11:36, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Thiago Braga Rodrigues**, **Usuário Externo**, em 02/03/2026, às 11:37, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Eduardo Lazaro Martins Naves**, **Professor(a) do Magistério Superior**, em 02/03/2026, às 11:37, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Ederson Rosa da Silva**, **Professor(a) do Magistério Superior**, em 02/03/2026, às 11:37, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Pierre Antoine Pino**, **Usuário Externo**, em 02/03/2026, às 11:37, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Guy Marc Alain Bourhis**, **Usuário Externo**, em 02/03/2026, às 11:37, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **7099953** e o código CRC **1AF246DE**.

Referência: Processo nº 23117.012493/2026-66 SEI nº 7099953

DEDICATION

To my grandmother, who taught me the value of education.

To my parents, for support and encouragement during difficult times.

ACKNOWLEDGEMENTS

This work would not have been possible without the help of my advisor, Prof. Dr. Eduardo Naves, for his constant guidance, patience, encouragement, valuable advice, career opportunities, and financial support.

I also thank MSc. Caroline Valentini, the healthcare professional who organized the tests with the patients and accompanied me in data collection, even during her pregnancy.

Thanks to all the complementary professionals who helped gather volunteers, as well as to the volunteers themselves, some with mobility difficulties, that nevertheless patiently attended the tests.

Thanks to the Coordination of Superior Level Staff Improvement (Capes) and the Coordination of the Faculty of Electrical Engineering for the scholarship, payment of the article processing fees and all support provided.

Finally, thanks to the journals and revisors that accepted my papers and to the examining board for their valuable time and effort.

ABSTRACT

Introduction: Stroke is a leading cause of long-term disability worldwide and frequently results in persistent upper-limb motor impairments. Accurate motor assessment is essential for rehabilitation planning and monitoring recovery. This assessment is traditionally performed using standardized clinical scales, such as the Fugl-Meyer Assessment (FMA), which require trained healthcare professionals, can be time- and resource-intensive, and involve subjectivity, as scores may vary across evaluators. Moreover, they cannot be easily administered remotely, and repeated in-person evaluations can be difficult for individuals with mobility limitations. Faster and more objective assessment approaches could enable more frequent evaluations and improve the tracking of motor recovery. Although recent advances in digital health, exergames, and Artificial Intelligence (AI) have enabled alternative solutions, many existing systems rely on specialized motion capture sensors, time-consuming software, or non-interpretable AI models, limiting accessibility, scalability, and clinical adoption. **Objectives:** This thesis aims to develop and evaluate a sensor-free, lightweight, and interpretable approach for post-stroke upper-limb motor assessment embedded into a rehabilitation exergame. The proposed system simultaneously provides therapeutic exercises and automatically estimates upper-limb motor performance during gameplay using only a standard camera, enabling supervised remote use. Specifically, the objectives are to (i) review existing automated motor assessment methods and their relationships with kinematic game parameters, and (ii) propose and preliminarily validate a low-cost exergame framework capable of estimating clinical motor impairment using simple and transparent kinematic features. **Materials and Methods:** A systematic review was conducted to analyze current technologies and computational approaches for automated upper-limb motor assessment after stroke, identifying the most used game parameters. Based on these findings, an experimental study was developed using a Unity-based rehabilitation exergame controlled by arm movements captured with a standard camera and processed using the MediaPipe framework. Sixteen kinematic and spatiotemporal features were extracted from two-dimensional hand and arm trajectories during gameplay. Twelve individuals with chronic stroke (24 upper limbs) participated, with bilateral FMA scores used as the clinical reference. Correlation analyses, exhaustive feature selection, and multiple linear regression modeling were performed to estimate FMA scores, with exploratory comparisons to alternative machine learning models. **Results:** The systematic review indicated that most existing approaches depend on external sensors or computationally complex methods, often limiting interpretability and real-world applicability. In the experimental study, some gameplay-derived and clinically interpretable features, including average hand aperture, and spatial exploration area, were significantly correlated with FMA scores. A lightweight multiple linear regression model demonstrated strong predictive performance for affected limbs (Spearman $\rho = 0.92$, $R^2 = 0.89$, RMSE = 4.42) and accurately stratified motor impairment severity, achieving accuracies between 86% and 93%. More complex machine learning models did not outperform the interpretable regression approach. **Conclusion:** This thesis demonstrates that sensor-free, low-cost, interpretable and lightweight computation models for post-stroke motor assessment can be embedded into rehabilitation exergames using kinematic features derived from a standard camera, and be potentially valid for motor function assessment. Further tests with more games and populations are needed to generalize the result. Integrating assessment into gameplay reduces clinical workload, enables high-frequency monitoring, and improves accessibility for telerehabilitation.

Keywords: Stroke rehabilitation; Motor assessment; Exergames; Telerehabilitation; Fugl-Meyer Assessment.

GRAPHICAL ABSTRACT

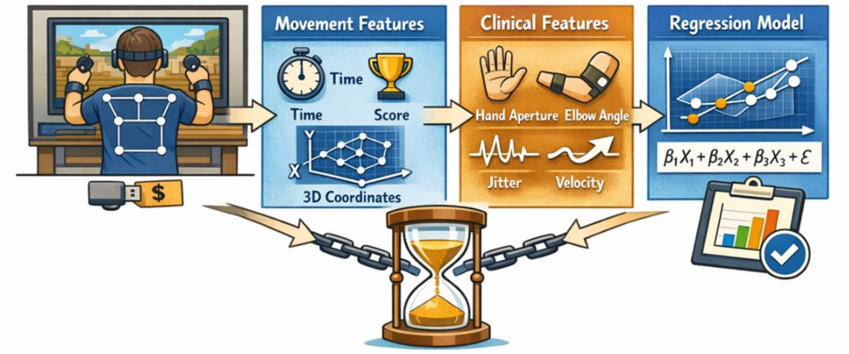
AI-DRIVEN LIGHTWEIGHT MOTOR ASSESSMENT EMBEDDED IN AN EXERGAME FOR POST-STROKE REHABILITATION

Stroke demographics

Traditional rehabilitation monitoring limitations

Hypothesis

“Clinically valid motor function can be estimated during gameplay (without separate evaluations) using only a camera-based exergame and clinically interpretable kinematic features”.

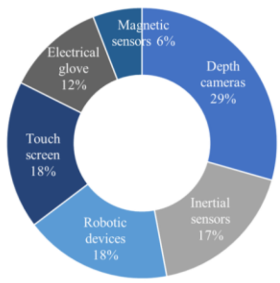


Phase 1: **Literature Review**

- Published in: **Medical & Biological Engineering & Computing, 2024 (IF 2.6)**.

- Studies that extracted game features and correlated them with any clinical scale;
- 14 studies;
- 244 participants.

Most used clinical scales found

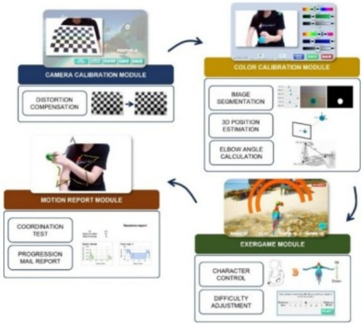


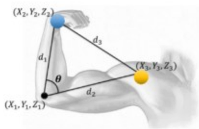
Most used devices found

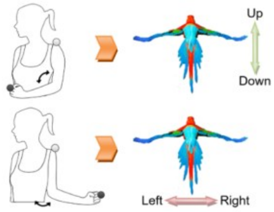
Correlation results (Game data vs clinical scores)

Phase 2: **Vision-Based Motion Tracking**

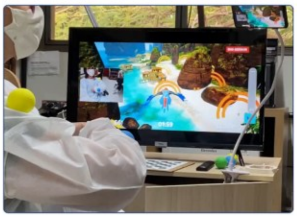
- Published in: **IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2025 (IF 5.2).**

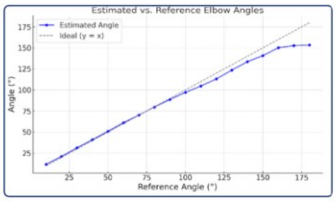


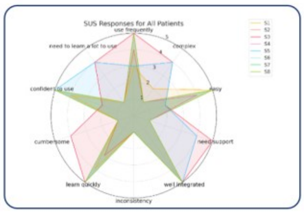




The first version of the system estimated elbow angles in real time using a standard camera



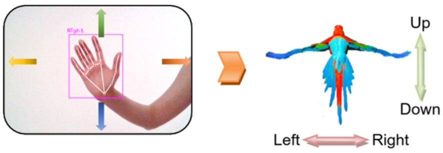




Preliminary test with 8 post-stroke subjects found an error in obtuse angles and excellent usability score

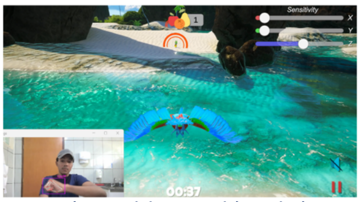
Phase 3: **AI-Driven Exergame with Embedded Assessment**

- Published in: **npj Digital Medicine, 2026 (IF 15.1)**.



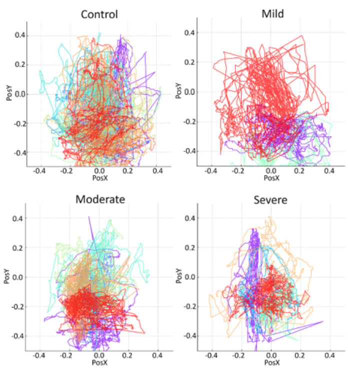
The game interaction method was changed to AI-based MediaPipe (Google). The wrist coordinates X and Y axes are used to move the character.

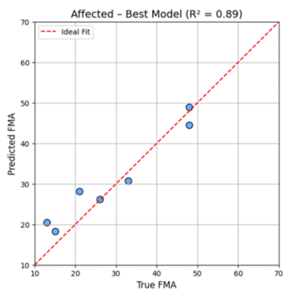
- 16 features were calculated from raw 3D coordinates, such as average hand angle, 2D area of movement (amplitude), trajectory dispersion and game score;
- A correlation analysis showed the most promising predictive features: average hand angle and correlation between shoulder and elbow angles.



12 stroke participants with varied ages and impairment severities

Wrist trajectory captured during gameplay, by motor severity, showing discriminative power of the features "2D Area" and "Total Traveled Distance", which measure amplitude of movement.





An exhaustive feature selection with 50/50 hold-out was conducted to find the equation for the multiple linear regression. This equation can predict the FMA score from game data.

Feature	State of the Art	Proposed Framework
Motion Capture	Depth cameras, IMUs, exoskeletons	Standard RGB camera only
Hardware Cost	Moderate to high	Low-cost
Setup Complexity	Calibration & sensors required	Plug-and-play
Assessment Mode	Separate from therapy	Embedded in gameplay (assessment can be more frequent and reduce therapist workload)
Interpretability	Often black-box ML	Interpretable features + regression
Remote Use	Limited	Designed for telerehabilitation
Scalability	Restricted	High
Computational Demand	Moderate to high	Lightweight

Scientific contributions of the proposed framework

Limitations:

- Small sample;
- FMA burden;
- The equation is game-specific;
- Flawed 3D depth estimation from MediaPipe.

Future work:

- Larger dataset;
- Other neurological conditions;
- Clinical scale easier to apply;
- Find definitive reliable measures that work among diverse game mechanics and populations;
- Integrate into telerehabilitation platform.

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
FMA	Fugl-Meyer Assessment
HMD	Head-Mounted Display
IMU	Inertial Measurement Unit
VR	Virtual Reality

TABLE OF CONTENTS

1 INTRODUCTION

The introduction clarifies the main concepts for understanding this work, its relevance, its context within the current literature, as well as its hypothesis and objectives.

1.1 Theoretical foundation

1.1.1 Stroke demographics

A cerebrovascular accident, commonly referred to as a *stroke*, is a major neurological condition and remains one of the leading causes of adult neurological impairment and long-term disability worldwide. A stroke affects around 795,000 individuals annually in the United States (2021), with overall stroke prevalence of 3.3% of the population (2017-2020) and increasing with advancing age in both males and females (MARTIN *et al.*, 2024). Globally, the burden of stroke continues to increase, with more than 93.8 million stroke survivors in 2021 (FEIGIN *et al.*, 2024). Among survivors aged 65 years and older, over half experience persistent reductions in mobility and motor control, which are strongly associated with a decline in quality of life (LOHSE *et al.*, 2014) and functional abilities (KIPER *et al.*, 2014).

1.1.2 Virtual Reality

Current theoretical frameworks define *Virtual Reality (VR)* as a computer-generated simulation that enables users to perceive and interact within a digital space as though it were real (MÉNDEZ *et al.*, 2025; SLATER; SANCHEZ-VIVES, 2016). Although VR has traditionally been associated with head-mounted displays (HMDs), such as the Apple Vision Pro (**Figure 1**), it can also be experienced through more accessible configurations, including smartphone or computer screens, which represent non-immersive or semi-immersive forms of VR (BOWMAN; MCMAHAN, 2007). These systems still provide meaningful virtual experiences by enabling real-time interaction and perceptual engagement, and align with the objective of this work, which is to test an accessible system. The adoption of HMD-based systems is limited due to their high cost, technical complexity, motion sickness concerns, and limited availability in institutional or clinical settings, making non-immersive solutions more feasible for broader implementation (ANTHES *et al.*, 2016; CIPRESSO *et al.*, 2018).

Figure 1 – Example of a Head-Mounted Display (Apple Vision Pro)

Source: VAUGHN RIDLEY / COLLISION / SPORTSFILE (2024). *Collision 2024 – VR3_1509*. Available at: https://commons.wikimedia.org/wiki/File:Collision_2024_-VR3_1509(53800584710).jpg.
Licensed under *Creative Commons Attribution 2.0 Generic (CC BY 2.0)*.

1.1.3 Serious games and exergames

Within this context, *serious games* are digital games designed with a primary purpose beyond entertainment, including education, training, health promotion, or rehabilitation. (MICHAEL; CHEN, 2006). Moreover, *exergames* represent a specific subset of serious games in which physical exercise constitutes a core interaction mechanism (MANSER *et al.*, 2025). An example of an exergame using motion-based interaction is shown in **Figure 2**, where a player performs physical gestures to control actions in a virtual golf game.

Figure 2 – Player interacting with a golf exergame with a television display (using Kinect motion sensor, pictured below the screen)

Source: SERGEY GALYONKIN (2013). *Kinect golf at Gamescom 2013*. Available at: https://commons.wikimedia.org/wiki/File:Kinect_golf_at_Gamescom_2013_(9588476123).jpg.
Licensed under *Creative Commons Attribution-ShareAlike 2.0 Generic (CC BY-SA 2.0)*.

1.1.4 Motion function rehabilitation and assessment

Upper limb rehabilitation in post-stroke patients typically involves structured, repetitive, and task-oriented exercises designed to promote neuroplasticity and restore voluntary motor control. Depending on severity and type of lesion, interventions may include range-of-motion exercises, constraint-induced movement therapy, robot-assisted training, and mirror therapy, which aim to stimulate motor pathways and improve coordination, strength, and dexterity (VEERBEEK *et al.*, 2014; WINSTEIN *et al.*, 2016).

Continuous monitoring of the patient's motor function is essential because stroke recovery is a dynamic and non-linear process influenced by neuroplasticity, therapy intensity, and adherence. Regular assessments allow clinicians to track progress, adjust rehabilitation goals, and personalize therapeutic interventions, ensuring that recovery strategies remain aligned with the patient's evolving abilities and limitations (BERNHARDT *et al.*, 2017; LANGHORNE; BERNHARDT; KWAKKEL, 2011).

In physiotherapy, there are methods for assessing motor functionality traditionally used and based on a broad scientific base for many decades, such as the Fugl-Meyer Assessment (FMA) (FUGL-MEYER *et al.*, 1975), the Wolf Motor Function Test (WOLF *et al.*, 2005) and the Box and Block Test (MATHIOWETZ *et al.*, 1985).

1.1.5 The Fugl-Meyer Assessment

The FMA, published in 1975, is among the most widely used and validated instruments available today, being considered one of the gold standards for post-stroke motor evaluation. It is a comprehensive and well-established assessment, noted for its reliability, sensitivity to change, and clinical applicability (GLADSTONE; DANELLS; BLACK, 2002).

This scale has been consistently reported in numerous studies in the current stroke rehabilitation literature using exergames (ALLEGUE *et al.*, 2022; BURDEA *et al.*, 2022; JIANG *et al.*, 2023). It consists of a checklist-based test, administered exclusively by a trained clinician who observes and scores the patient's ability to perform specific motor tasks, such as reflexes, flexion, extension, and coordination movements of the elbow, shoulder, and hand, as well as their synergistic patterns. Each movement is rated on a 3-point ordinal scale (0 = cannot perform, 1 = performs partially, 2 = performs fully). The upper-limb section of the FMA is a subsection of the full test, comprising 33 items and a maximum score of 66 points, which represents normal motor function. The FMA was used in this preliminary study as the gold-standard comparison method. The complete version of the FMA used in this study is presented in **Annex I**.

Although the FMA is widely recognized for its comprehensiveness and psychometric robustness, its clinical application can be challenging in practice. The test cannot be performed remotely because it requires specific clinical instruments, while many post-stroke individuals present mobility limitations. It also requires a trained examiner, can take considerable time to complete: often between 30 and 60 minutes, and involves subjective scoring that may vary among evaluators. In addition, performing all test items can be difficult for individuals with severe cognitive impairments or limited endurance, occasionally resulting in incomplete assessments (GLADSTONE; DANELLS; BLACK, 2002; PAGE; LEVINE; HADE, 2012).

These limitations can hinder its use both in regular rehabilitation with a clinician or in remote monitoring contexts, emphasizing the need for simplified, automated, or technology-assisted alternatives that preserve the validity of motor performance evaluation while reducing examiner dependency and logistic burden.

1.1.6 Clinical applications of Virtual Reality and opportunities

In recent years, VR has become a multidisciplinary tool utilized in clinical medicine for a variety of purposes, such as pain management (POURMAND *et al.*, 2018), neurocognitive impairment assessment (YEH *et al.*, 2012), medical skill teaching (BARTEIT *et al.*, 2021), and physical rehabilitation (POURMAND et al., 2017; SÁNCHEZ-HERRERA-BAEZA et al., 2020). The scientific literature indicates that VR has applications for visual, auditory, tactile, and motor learning; it has a favorable impact on self-motivation; and it has also been used to enhance post-stroke motor skills (PARK; LEE; LEE, 2013). Regarding this last area of research, VR has not only been used for post-stroke rehabilitation (AHMAD *et al.*, 2019; NOROUZI-GHEIDARI *et al.*, 2020; OLIVEIRA *et al.*, 2026; WEBER *et al.*, 2019), but also shows potential for assessing motor functionality (ADAMS *et al.*, 2019; KIM *et al.*, 2016; MASMOUDI *et al.*, 2024; RODRIGUEZ-DE-PABLO, C *et al.*, 2015).

As VR systems become increasingly capable of replicating real-world physical environments, they present new opportunities for automated, remote, and continuous motor assessment (LEE; LEE; KIM, 2018). VR-based training sessions can potentially reproduce established clinical tests or correlate gameplay-derived parameters with standardized evaluation metrics, such as the FMA (ADAMS *et al.*, 2019; RAHMAN *et al.*, 2023; TANNUS; NAVES; MORERE, 2024). However, such integration of digital assessment remains limited in current clinical practice, primarily due to the lack of standardized protocols, interoperability challenges, and accessibility constraints (MAGGIO *et al.*, 2025). This gap highlights a promising area for further research and development, especially with the emergence of low-cost, vision-based and AI-assisted rehabilitation systems (KLEIN *et al.*, 2024; RAHMAN *et al.*, 2023).

Also, it would be interesting to further increase the independence of virtual therapy from the real world, through telerehabilitation, as virtual therapy can offer opportunities to enhance flexibility, convenience, cost-effectiveness, personalization, engagement, remote monitoring, accessibility, data collection, and progress monitoring (CARMONA *et al.*, 2023; POURMAND *et al.*, 2017b; RODRIGUEZ-DE-PABLO, CRISTINA *et al.*, 2012).

1.1.7 Motion tracking for exergames

In exergame-based rehabilitation and monitoring, accurate motion tracking is essential to ensure that a patient's movements are correctly recognized, measured, and translated into interactive virtual tasks. For post-stroke individuals, especially those with upper-limb impairments, motion tracking must accommodate limited and asymmetric ranges of motion, reduced fine-motor control, and compensatory movement patterns (TANNUS; NAVES; MORERE, 2024; THOMSON *et al.*, 2020). Various technologies have been used for this purpose, including optical tracking systems (such as Microsoft Kinect and infrared cameras), inertial measurement units (IMUs), wearable sensors, and electromechanical devices such as exoskeletons or robotic arms (FABBRIZIO *et al.*, 2023; GU *et al.*, 2022; TANNUS; NAVES; DE SÁ, 2021).

Although these systems can provide high-precision kinematic data, they also present significant challenges for clinical and home-based applications. Optical systems typically require specialized hardware, dedicated space, and proper lighting conditions. Wearable sensors and exoskeletons, while capable of precise motion capture, are expensive, cumbersome to calibrate, and difficult to use for patients with limited mobility or spasticity, often requiring therapist assistance, which limits their accessibility outside laboratory environments (THOMSON *et al.*, 2020). Furthermore, the maintenance and calibration demands of such equipment increase costs and reduce scalability in low-resource healthcare settings (THOMSON *et al.*, 2016).

In recent years, advances in computer vision and artificial intelligence (AI) have enabled new approaches to human motion tracking without physical markers or specialized hardware. AI-based pose estimation models can analyze standard video input from regular cameras to identify body landmarks and estimate joint angles in real time (LUGARESI *et al.*, 2019). These methods significantly reduce cost and increase accessibility, opening new possibilities for remote, low-cost, and scalable rehabilitation systems that can be integrated into exergames and home-based motor assessment platforms (FABBRIZIO *et al.*, 2023; GU *et al.*, 2022; RAHMAN *et al.*, 2023).

1.1.8 MediaPipe

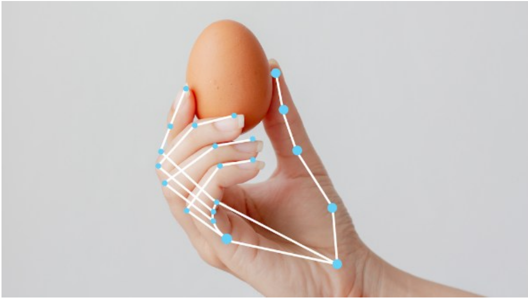
MediaPipe (GOOGLE, 2026), an open-source framework developed by Google Research, represents a major milestone in AI-based human motion tracking. Initially released in 2019, with the full-body solution introduced in 2021, it enabled a paradigm shift in real-time pose estimation by optimizing the processing architecture of deep learning models (LUGARESI *et al.*, 2019). Before its introduction, most AI-based pose estimation frameworks were computationally demanding, typically requiring dedicated GPUs or high-end workstations, which limited their use on consumer-grade devices (CAO *et al.*, 2021; NEWELL; YANG; DENG, 2016). In contrast, MediaPipe enables the detection and tracking of multiple human joints directly on low-power devices, mobile platforms, and web browsers, achieving substantially higher efficiency than earlier systems.

Beyond its relatively lightweight computational design, a key advantage of MediaPipe is its independence from depth sensors, external markers, or specialized motion capture equipment. This characteristic makes the framework particularly suitable for low-cost, scalable, and remote rehabilitation applications, where accessibility and ease of deployment are essential.

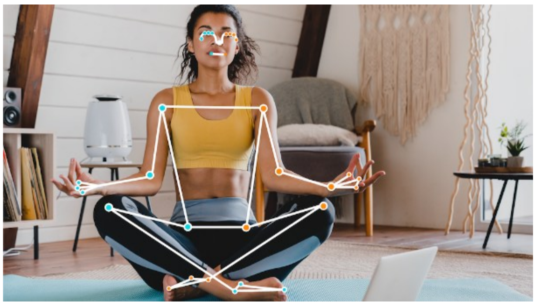
In this work, two models from the MediaPipe framework were employed: the full-body and hand estimation models, illustrated in **Figure 3**. Both models estimate anatomical landmarks using only a standard color camera and provide 3D coordinates for each landmark at every frame.

Despite the advantages, MediaPipe presents limitations when compared to depth-sensing systems such as Microsoft Kinect. As it relies exclusively on 2D camera input, its 3D reconstruction is inference-based rather than derived from direct depth measurements, meaning that the Z-coordinate does not represent absolute depth. Consequently, estimation errors tend to increase in scenarios involving occlusions, rapid movements, or complex joint rotations (GUO *et al.*, 2025).

Figure 3 – MediaPipe hand and full-body landmark estimation models



(a)



(b)

(a) Hand landmark estimation model, illustrating the detection of anatomical key points of the hand using a monocular RGB camera.
(b) Full-body pose landmark estimation model, showing upper- and lower-body joint tracking from RGB input.
Source: *MediaPipe documentation* (GOOGLE, 2026).
Licensed under *Creative Commons Attribution 4.0 International (CC BY 4.0)*.

Nevertheless, MediaPipe can provide a favorable balance between accuracy, portability, accessibility, and usability that remains unmatched by traditional hardware-dependent motion capture systems.

1.2 State of the art

1.2.1 Related works

In the context of the emerging technological opportunities, in recent years, some initiatives have sought to automate post-stroke motor assessment through digital, sensor-based, or AI-assisted approaches. These studies vary widely in terms of sensing modality, computational complexity, and clinical applicability.

Some approaches focused on mobile-based evaluations. For instance, SONG; CHEN; et al., (2019) developed a smartphone-based system that estimated upper-limb motor scores using the smartphone's embedded inertial sensors and decision tree algorithms, achieving significant correlation with the Fugl-Meyer Assessment (FMA) ($R^2 = 0.78$).

Other studies incorporated commercial motion sensors or optical systems to increase accuracy. NIE et al., (2021) proposed an open-source portable solution to estimate wrist kinematics from two commercial devices, one using IMUs and another using optical markers. BAI et al., (2015) also utilized IMUs to quantify upper-limb motion during neurorehabilitation, demonstrating that wearable sensors can provide precise temporal and spatial data correlated with motor recovery. PALANI et al., (2022) combined the MediaPipe framework with IMUs to estimate joint angles in real time, validating the hybrid potential of combining computer vision with wearable technologies. Likewise, CHEN et al., (2023) proposed an automatic quantitative assessment tool that predicts the hand sub-score of the FMA scale using special optical motion capture devices' data and decision trees, achieving an RMSE of 17.4 points in a dataset of 79 participants. Although effective, these systems required external hardware and calibration, limiting their scalability and user independence. Also, they were not integrated into a rehabilitation software.

Other recent advances aimed to integrate the automated motor assessment into the rehabilitation game, although not eliminating physical sensors. JIANG et al., (2023) introduced a Kinect-based serious game that applies fuzzy inference to estimate motor function, obtaining an accuracy of 93.5% when compared with the FMA.

In a comprehensive review, RAHMAN et al., (2023) analyzed AI-driven rehabilitation and assessment systems, concluding that although deep learning models can achieve high predictive accuracy, they generally lack interpretability and demand large, annotated datasets. The need for explainable, low-complexity solutions remains evident.

Finally, more interactive approaches have integrated rehabilitation and assessment into serious games. SONG; DING; et al., (2019) also explored augmented-reality games to enhance engagement and mental state while tracking motion recovery. JIANG et al., (2023)'s serious game demonstrated that gamified assessment can both motivate patients and objectively quantify motor function. These findings support the integration of functional evaluation into therapeutic gameplay, a concept further advanced by the present work.

1.2.2 Challenges and research opportunities

While many of these systems are promising, they often require dedicated assessment time, lack interpretability, need external sensors (such as depth cameras or IMUs), rely on complex architectures, and cannot be applied remotely. In contrast, combining motor assessment directly into the rehabilitation game using only the game display device camera could offer potential advantages: it could decrease the need for separate clinical evaluations, reduce monotony, increase engagement, and enable real-time, high-frequency tracking of recovery without burdening therapists or patients. This integration could promote scalability, remote use, and personalization, allowing stroke survivors to be continuously monitored during gameplay using simple, transparent metrics that could reflect functional performance, facilitate clinical understanding, and support generalization to other exergames in the future.

1.3 Hypothesis

The hypothesis explored in this work is that clinically meaningful assessment of post-stroke upper-limb motor function can be achieved directly during therapeutic gameplay, using only a low-cost, camera-based exergame powered by AI motion capture (MediaPipe). Specifically, it proposes that simple and interpretable kinematic features extracted in real time from 3D-simulated hand and arm coordinates, such as hand aperture, trajectory, range of motion, and movement smoothness, can accurately predict standardized clinical scores (e.g., the FMA) without the need for specialized hardware or separate evaluation sessions.

1.4 Objectives

1.4.1 General objective

This research aims to develop and preliminarily validate a VR rehabilitation exergame capable of simultaneously providing therapeutic engagement and quantitative motor assessment for post-stroke patients. The study seeks to determine whether kinematic and gameplay-derived features obtained during therapeutic sessions can be used to estimate the score of a standardized clinical scale (the FMA), thereby demonstrating the potential of VR game data as an indicative of motor recovery for remote monitoring.

1.4.2 Specific objectives

To achieve this goal, the project was structured into three progressive phases, each addressing distinct methodological and developmental aspects of the system.

Phase 1: Literature review

To conduct a comprehensive literature review to examine existing research on:

- VR-based upper-limb rehabilitation in stroke patients;
- Game-based assessments for post-stroke motor function;
- Correlations between VR game data (e.g., 3D coordinates, game score, and time) and established clinical scales for upper-limb function assessment (e.g., FMA);
- Feasibility and usability of VR systems for home-based rehabilitation in stroke patients.

Phase 2: Pilot study with therapist

- Refine the design of an exergame for clinical relevance and effectiveness;
- Define specific VR game features for data collection;
- Pilot test the VR system with post-stroke patients under therapist supervision, collecting gameplay data and clinical scale scores;
- Analyze the correlation between VR game data and clinical scale scores to assess the potential validity of the VR system as an assessment tool.

Phase 3: Preliminary validation

- Further improve the interaction and design of the exergame;
- Retest the system with more post-stroke patients;
- Extract interpretable kinematic and spatiotemporal features (e.g., hand angle, movement area, and trajectory) and correlate them with clinical FMA scores;
- Evaluate the predictive capacity of lightweight statistical models (e.g., linear regression) to estimate FMA-equivalent digital scores from gameplay data.

1.5 Justification and relevance

1.5.1 Reduction of evaluator subjectivity in motor assessment

A central limitation of conventional clinical scales used in post-stroke rehabilitation is their dependence on the evaluator's judgment. Although the Fugl-Meyer Assessment is extensively validated and widely adopted, its scoring relies on visual inspection and manual rating by trained examiners, which may introduce inter-rater and intra-rater variability (DUNCAN; PROPST; NELSON, 1983; GLADSTONE; DANELLS; BLACK, 2002). By extracting quantitative kinematic features directly from patient movement, the proposed approach can reduce reliance on subjective interpretation and introduce a more objective layer of motor assessment. This can contribute to greater consistency in monitoring motor performance across sessions, evaluators, and clinical contexts.

1.5.2 Patient motivation and engagement through playful assessment

Traditional motor assessments are often perceived by patients as time-consuming and repetitive, which may negatively affect engagement, particularly during long-term rehabilitation (GLADSTONE; DANELLS; BLACK, 2002). Integrating evaluation into the rehabilitation exergame can transform assessment into a faster and more enjoyable activity, embedded within a playful and goal-oriented task. This design can increase patient motivation and adherence, as the evaluation process becomes part of an engaging therapeutic experience rather than a separate clinical procedure.

1.5.3 Remote assessment and rehabilitation feasibility

The proposed framework can enable both rehabilitation and motor assessment to be performed remotely, using commonly available devices, such as smartphone, tablet or computer. This capability can be especially relevant for post-stroke individuals who face transportation difficulties or geographic barriers to accessing clinical facilities. Remote assessment also can benefit therapists by reducing the need for frequent in-person evaluations, enabling a better time efficiency in their appointments, as well as supporting more flexible and scalable rehabilitation workflows.

1.5.4 Cost reduction for patients and healthcare systems

Another important justification for this work is the potential reduction in economic burden associated with post-stroke rehabilitation using games. By eliminating the need for specialized hardware, such as depth cameras, wearable sensors, or robotic systems, the proposed solution can lower equipment costs. Additionally, partial automation of the assessment process reduces the number of therapist hours dedicated exclusively to evaluation, while remote use minimizes patient transportation expenses. Together, these factors contribute to a more cost-effective and sustainable rehabilitation model.

1.5.5 Interpretability and generalization of assessment outcomes

Unlike computational algorithms that provide limited insight into how predictions are generated, this work emphasizes the use of simple and interpretable kinematic features (e.g., average hand aperture, maximum 2D movement range). These features are directly linked to observable motor behaviors, facilitating clinical understanding and trust in the assessment outcomes. Importantly, this interpretability allows the framework to be generalized to other exergames with similar interaction mechanics and can support future game design decisions. Even if the exact estimation of standardized clinical scores is not sufficiently validated or trusted, the extracted metrics can still provide an approximation of motor status and recovery trends, enabling continuous monitoring and supporting data-driven therapeutic decisions.

1.6 Chapter organization

This doctoral thesis was written in an alternative article-based format, in accordance with the regulation of the Postgraduate Program. Therefore, this work is organized as follows:

Chapter 1: Introduction

Chapter 1 provides the conceptual and theoretical background necessary to understand the research, including its context, relevance, and objectives. The opening sections present the theoretical foundation on stroke rehabilitation, VR, exergames, motion tracking technologies, and AI-based assessment systems. Then, the state of the art is analyzed, finding potential research opportunities. Accordingly, the hypothesis and objectives of this study are defined and justified. Finally, the current section outlines the overall structure of the thesis.

Chapter 2: Results and discussion

This chapter presents an integrated discussion of the results and scientific contributions from the three peer-reviewed papers presented in this document. It analyzes their methodological evolution, the relationship between their findings, and their impact on post-stroke rehabilitation.

The first paper, *Post-Stroke Functional Assessments Based on Rehabilitation Games and Their Correlation with Clinical Scales: A Scoping Review* (Medical & Biological Engineering & Computing) (TANNUS; NAVES; MORERE, 2024) corresponds to Phase 1 of the study objectives, and reviewed the correlation between rehabilitation game metrics and clinical scales, emphasizing that interpretable models embedded in gameplay can provide reliable digital biomarkers and reduce therapist workload.

The second paper, *Low-Cost Vision-Based 3-D Elbow Tracking for Post-Stroke Rehabilitation: Development and Pilot Evaluation of a Serious Game* (IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2025) (TANNUS *et al.*, 2025), corresponds to Phase 2 of the study objectives, describing the design and pilot testing of a low-cost exergame using RGB-based 3D elbow tracking.

The third paper, *AI-Driven Low-Cost Rehabilitation Exergame as a Lightweight Framework for Stroke Assessment* (TANNÚS; VALENTINI; NAVES, 2026), corresponds to Phase 3, introducing an AI-powered exergame that integrates automated motor assessment during gameplay.

Chapter 3: Conclusion

A conclusion regarding all the contributions of this thesis is presented, as well as future works. The final subsection lists all the conference and journal productions published during the doctorate studies.

Annexes

The annex compiles supplementary materials relevant to the thesis, including Annex I: Full version of the Fugl-Meyer Assessment for Upper Limbs.

2 RESULTS AND DISCUSSION

This chapter discusses the main findings and scientific contributions of the three published works derived from this doctoral research, highlighting their methodological evolution, interconnection, and impact on the field of technology-assisted stroke rehabilitation. Together, these studies illustrate a progressive research trajectory from the development of a low-cost, vision-based exergame for motor training to the creation of an AI-driven framework potentially capable of automated functional assessment during gameplay.

2.1 Literature review

2.1.1 Context and Summary

This section presents the full reproduction of the article entitled *Post-stroke functional assessments based on rehabilitation games and their correlation with clinical scales: A scoping review*, published in Medical & Biological Engineering & Computing (TANNUS; NAVES; MORERE, 2024). The reproduction of this material has been formally authorized by Springer Nature, in accordance with the publisher's copyright policies.

This article constitutes a fundamental component of the present doctoral thesis, corresponding to Phase 1 of the research objectives, which focuses on the systematic investigation of existing approaches for automated motor assessment in post-stroke rehabilitation using VR and game-based systems. It establishes the theoretical and methodological foundation upon which the subsequent experimental developments of this thesis are built.

The primary objective of the study was to evaluate whether motion data automatically collected during interaction with rehabilitation games and VR simulations can be considered clinically valid when compared to standardized physiotherapy assessment scales. To this end, a scoping review was conducted following PRISMA guidelines, analyzing studies published over the past decade that correlate in-game parameters with conventional clinical measures of upper-limb motor function.

A total of 14 studies involving 244 post-stroke participants were included in the review. The findings indicate that 85.7% of the studies reported positive or strongly positive correlations between game-derived metrics and clinical scales such as the FMA, Wolf Motor Function Test, Box and Blocks Test, and Action Research Arm Test. The most used in-game parameters were categorized into three main groups: (i) game scores, (ii) kinematic variables (e.g., movement speed, smoothness, joint angles), and (iii) time-based metrics.

Despite these promising results, the review also highlights important limitations in the current state of the art. Most existing systems rely on specialized hardware, such as motion capture sensors, robotic devices, or depth cameras, which restrict accessibility, scalability, and real-world applicability. Additionally, the heterogeneity of methodologies and the relatively small sample sizes limit the generalizability of the findings.

Therefore, this article provides critical evidence supporting the feasibility of using game-derived data as digital biomarkers for motor function assessment, while simultaneously identifying key research gaps. These gaps directly motivate the contributions of this thesis, which aims to develop a low-cost, sensor-free, and interpretable assessment framework embedded within an exergame, using only standard camera input and lightweight computational models.

Medical & Biological Engineering & Computing (2024) 62:47–60
https://doi.org/10.1007/s11517-023-02933-9

REVIEW ARTICLE

Post-stroke functional assessments based on rehabilitation games and their correlation with clinical scales: A scoping review

Julia Tannus[1] · Eduardo L. M. Naves[1] · Yann Morere[2]

Received: 28 March 2023 / Accepted: 7 September 2023 / Published online: 19 September 2023
© International Federation for Medical and Biological Engineering 2023

Abstract

Considering that stroke is one of the main causes of adult impairment and the growing interest in Virtual Reality (VR) as a potential assessment and treatment tool for the rehabilitation of stroke patients, a scoping review was conducted to check whether user's motion data obtained from VR games and simulations can be clinically valid. This was done by reviewing studies on parameters for assessing the functional skills and rehabilitation progress using data from VR games or simulations. Then, identifying the most widely used and validated parameters for the quantification of motor ability in a virtual environment and suggesting challenges for future research. For the validation of the parameters obtained from the VR software, only the studies that correlated them with traditional physiotherapy scales were considered. In December 2022, a search of the following databases was performed: IEEE Xplore, ACM Digital Library, PubMed and PEDro. The selection criteria were studies published in English during the past 10 years, with upper-limb based interaction and tested on more than one stroke patient. A total of 14 were included in the PRISMA scoping review. Favorable results were found in 12 of the 14 studies, which reported positive or strongly positive correlations with clinical scales, even when diverse variables were used. In-depth research using a larger sample size is needed. The results demonstrate that data collected while playing a virtual serious game has the potential to be clinically valid, after conducting high-quality supportive studies with controlled variables, potentially helping the practice in terms of time and resources.

Keywords Virtual reality · Stroke · Rehabilitation · Scoping review · Games

Abbreviations

ARAT	Action Research Arm Test
BBT	Box and Blocks Test
CAHAI	Chedoke Arm and Hand Activity Inventory
FMA	Fugl-Meyer Assessment
MAL	Motor Activity Log
MAS	Modified Ashworth Scale
MMSE	Mini-Mental State Examination
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
TEMPA	Upper Extremity Performance Test for the Elderly
VR	Virtual Reality
WMFT	Wolf Motor Function Test

✉ Julia Tannus
julia.tannus95@gmail.com

1 Faculty of Electrical Engineering, Assistive Technologies Group, Federal University of Uberlandia, Av Joao Naves de Avila, 2121 –Bloco 1A, Uberlandia, Brazil

2 LCOMS Laboratory, University of Lorraine, Metz, France

1 Introduction

A cerebrovascular accident, the medical term for what is commonly known as stroke, is a type of neurological disorder and the main reason for adult neurological impairment and serious long-term disability. Moreover, every 40 s, someone in the USA has a stroke, totaling more than 795,000 people every year [1]. In more than half of stroke survivors above age 65, mobility remains reduced. These motor abnormalities are linked to a decline in quality of life [2] and functional abilities [3].

To restore neurological functions that have been lost or impaired as a result of stroke brain damage, a variety of techniques can be used [4]. While Virtual Reality (VR) first gained popularity in the late 1980s and early 1990s, it has only just began to be developed and explored during the past ten years as a potential tool for assessment and treatment within the neurorehabilitation field [5]. The cost of the devices has drastically decreased, and widespread access to high-speed internet connection, together with technological advancements in the systems, have increased their use [6].

In several technological fields, particularly in videogames, VR has recreated itself. On various commercial platforms, including PlayStation, these kinds of games have gained popularity.

VR is the use of a user-computer interface that simulates, in real-time, an area, scenario, or activity, allowing the user to interact through many sensory inputs [1, 7]. In recent years, VR has become a multidisciplinary tool utilized in clinical medicine for a variety of purposes, such as pain management [8], neurocognitive impairment assessment [9], medical skill teaching [10], and physical rehabilitation [11]. The scientific literature indicates that VR has applications for visual, auditory, tactile, and motor learning; it has a favorable impact on self-motivation; and it has also been used to enhance post-stroke motor skills [12]. Regarding this last area of research, VR has not only been used for post-stroke rehabilitation [7, 13, 14], but also shows potential for assessing motor functionality [15–17].

VR games and simulators have several classification subcategories. One of the important distinctions is related to the user's level of immersion, which can be classified into two groups: immersive and non-immersive systems [18]. Immersive VR systems can use, for instance, VR glasses, projection on a large screen or in caves, leading the user to experience a disconnection with the real world. On the other hand, non-immersive VR systems are those used on common computer screens, video game consoles and mobile devices, in which the user can still visualize the real world; however, his mental presence is focused on the virtual simulation. In addition to the display, there are also other input and output devices that can be used in VR, for example, motion trackers, exoskeletons, sensors, bionic gloves, treadmills, etc. These systems are interesting resources to produce an environment similar to the real one, in which it is possible to manipulate the content, duration, intensity and feedback to create a prescription for physical exercises at home. Training management is related to the quality of recovery after stroke. Moreover, the main advantage of VR interventions over conventional treatment is that patients see it as a fun exercise game rather than a monotonous treatment, which can be used anytime and anywhere, leading to improvements in motivation and adherence to treatment [19].

Considering the current literature, more than a decade ago, in 2007, Henderson et al. [18] concluded that the scientific evidence on VR in stroke rehabilitation was limited. Later, in 2014, Lohse et al. [2] stated that VR therapy represents a significant advantage over conventional therapy in terms of body function, activity results, and participation in the treatment. These authors also stated that research in this field was limited. More recently, in 2017, a Cochrane review performed by Laver et al. [20] showed favorable results of VR in addition to conventional therapy on upper limb motor function, but that VR was comparable to conventional therapy. This implies that more evidence can be collected to affirm the effectiveness of VR therapy.

Also, it would be interesting to further increase the independence of virtual therapy from the real world, as the virtual therapy can offer opportunities to enhance the flexibility, convenience, cost-effectiveness, personalization, engagement, remote monitoring, accessibility, data collection and progress monitoring. In traditional physiotherapy, there are methods for assessing motor functionality traditionally used and based on a broad scientific base for many decades, such as the Fugl-Meyer Assessment [21], the Wolf Motor Function Test [22] and the Box and Block Test [23]. As VR is increasingly capable of simulating the real world, the patient potentially could also be evaluated from data collected while playing VR games, either by reproducing the test in the virtual world or capturing parameters and correlating them with the conventional methods. This, however, was not found as being a common practice today. In this way, the purpose of this scoping review was to identify initiatives that aimed to collect real-time parameters during gaming or simulations and correlate them with conventional measures of motor skill. The review examined the scientific evidence generated and explored the potential of this method.

2 Materials and methods

2.1 Search strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards for systematic reviews and meta-analyses were followed when conducting and reporting the current review [24].

A search of the literature was made in December 2022, on the following databases: IEEE Xplore, ACM Digital Library, PubMed and PEDro. The Boolean search terms utilized are described in Table 1.

2.2 Eligibility criteria

The eligibility criteria used to select the articles were: (i) published in the last 10 years (2012–2022), in order to get updated research information; (ii) English language; (iii) not

Table 1 Boolean search terms for the systematic review

Games	game
Stroke	AND stroke
Functional assessment	AND (function assessment)
Excluding factors: reviews	NOT ("meta-analysis" OR review)

reviews or meta-analyses; (iv) any study design; (v) tested on more than one adult stroke patient; (vi) intervention based on video games that use immersive, semi-immersive, and non-immersive VR technologies to simulate virtual environments on computers, video game consoles, mobile apps, and VR glasses; (vii) upper limbs-based interaction; (viii) computerized assessment parameters generated during-play; (ix) correlation of this parameters with a standard clinical scale for motor function of upper limbs or cognition. Studies whose full text was not accessible were excluded.

2.3 Study selection and data extraction

We started by conducting a search using the combination of Boolean terms across all the databases mentioned previously. After reading the title and abstract, potentially relevant articles were found, and duplicates were removed, using an automated tool. To gather the papers included in the current scoping review, a thorough verification of conformity with the inclusion criteria was then completed.

The process of choosing studies was made by two reviewers, carefully extracting data from each one. From each article in the review, the following details were taken out: authors, country, conference or journal, year, number of participants, intervention (the software description), interaction method (device or artifact used to play the game or simulation), in-game assessment parameters (the parameters automatically collected by the software while playing), clinical scales used for correlation (with the computerized parameters) and the correlation results obtained.

In addition, we considered the population, intervention, control, outcome, and study design (PICOS) approach [24] to design the study selection:

Population: VR games and simulations for post-stroke rehabilitation;
Intervention: any user test that involves playing the post-stroke game or simulation;
Control: none taken into account;
Outcome: user's data collected through the virtual game or simulation and the application of a traditional physiotherapy scale for assessing the upper limbs post-stroke;
Study design: Randomized Controlled Trial, cohort, case–control, case series, observational, quasi-experimental or crossover studies.

3 Data analysis

Some data of the studies were extracted to better characterize them. Tables and graphs were made to present the most relevant information. They are: frequency of studies by country; frequency of studies in journal or conference; frequency of studies by year and weighted average of years; box plot of the number of subjects per study, showing maximum, minimum, mean and overall distribution; most used clinical scales; and correlations outcomes distribution. The software used to create the figures and tables was Microsoft Excel.

4 Results

4.1 Search results

Database searches were performed using the above-mentioned combination of keywords, and a total of 340 documents were retrieved. After duplicates were eliminated, 327 articles were screened for eligibility. In total, 14 studies were selected to be reviewed and analyzed. Figure 1 describes in detail the entire selection process in the respective phases. The primary characteristics of the studies are displayed in Table 2.

4.2 General data

Regarding the countries from which the studies are, it was found that most of them came from the USA, with a total of seven studies. Greece, Australia, The Netherlands, Spain,

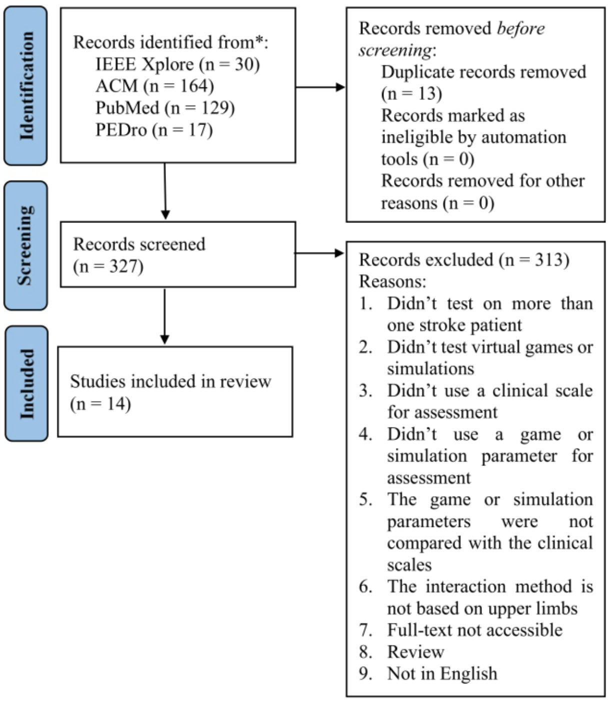


Fig. 1 PRISMA 2020 flow diagram for the studies selection [15]. Alt Text: A diagram with studies identified, screened and included, with reasons. 327 studies were screened and 14 were included

Table 2 Main characteristics of the study interventions

n	Study	Country	Conference or Journal	Year	Population	Intervention	Interaction method	In-game assessment parameters	Clinical scale used for correlation	Correlation result
1	Astrakas et al. [25]	Greece and USA	Journal	2021	N = 8	A labyrinth serious game	MR_CHIROD, a robotic device	Serious game score elements (collisions, rewards and total score)	FMA, ARAT, MAS and BBT	Positive
2	Lee et al. [26]	China	Journal	2021	N = 22	A basketball game	Kinect	Movement offset (root mean square), speed, efficiency, maximum instantaneous speed, average speed, V-variation, basketball holding time; neural networks	FMA, TEMPA and WMFT	Mixed
3	Jung et al. [27]	USA and Korea	Journal	2019	N = 12	Neuro-World, a collection of innovative 3D mobile games	Touch screen	Highest difficulty level reached; shortest time, longest time, and mean time to clear stages and to select answers; stages tried, cleared and failed; average and overall scores	MMSE	Positive

Table 2 (continued)

n	Study	Country	Conference or Journal	Year	Population	Intervention	Interaction method	In-game assessment parameters	Clinical scale used for correlation	Correlation result
4	Bai et al. [28]	China	Journal	2019	N = 32	HomeRehabMaster system, with 4 games	Kinect and inertial sensor	Elevation angle and azimuth angle of the shoulder joint relative to the center point of the shoulder; elevation angle and azimuth angle of the elbow joint relative to the shoulder joint; elevation angle and azimuth angle of the wrist relative to the elbow joint; angle of the forearm from posture sensor; distance between hand and head, and elbow and head; upper limb relative surface area from the reachable workspace	FMA	Positive
5	Jung et al. [29]	USA and Korea	Conference	2018	N = 12	Neuro-World, a collection of mobile games	Touch screen	Highest difficulty level reached; shortest time, longest time, and mean time to clear stages and to select answers; stages tried, cleared and failed; average and overall scores	MMSE	Positive
6	Lin et al. [30]	China	Journal	2018	N = 15	VR game	Inertial sensor	Game score	FMA	Positive
7	Hesam-Shariati et al. [31]	Australia	Journal	2017	N = 24	Wii-based Movement Therapy	Wii	Game score	WMFT, FMA and MAL	Positive

Table 2 (continued)

n	Study	Country	Conference or Journal	Year	Population	Intervention	Interaction method	In-game assessment parameters	Clinical scale used for correlation	Correlation result
8	Cidota et al. [32]	The Netherlands	Conference	2017	N = 10	An augmented reality game, called "post office trouble"	Kinect	Total duration, grasping (mean and standard deviation of the linear distance between the tip of the thumb and the index finger), max elbow angular velocity, max upper arm angular velocity, max wrist velocity, max trunk velocity, standard deviation of elbow angle, standard deviation of upper arm angle, wrist trajectory, trunk displacement	ARAT	Negative
9	Chen et al. [33]	China	Conference	2017	N = 21	VR rehabilitation games: basketball throwing; ball juggling; and window wiping	Kinect	Mean speed, mean velocity, smoothness	FMA and WMFT	Positive
10	Rodriguez-de-Pablo et al. [17]	Spain and Serbia	Conference	2015	N = 19	Serious games	A low-cost robotic system	Range of movement, control of movement, range of force, control of force	FMA, ARAT and WMFT	Positive
11	Friedman et al. [34]	USA	Journal	2014	N = 12	Frets on Fire game	MusicGlove	Game score	BBT	Positive
12	Khademi et al. [35]	USA	Conference	2014	N = 14	Modified Fruit Ninja game	Leap Motion	Game score	FMA and BBT	Positive

Table 2 (continued)

n	Study	Country	Conference or Journal	Year	Population	Intervention	Interaction method	In-game assessment parameters	Clinical scale used for correlation	Correlation result
13	Serradilla et al. [36]	UK	Conference	2014	N = 33	Circus Challenge game	Magnetic motion tracking	Position and orientation data are processed and 320 kinematic variables are derived from it. Those variables are related to underlying features which define how well the movement is performed; namely, speed, fluency, synchrony and accuracy	CAHAI	Positive
14	Friedman et al. [37]	USA	Conference	2011	N = 10	Frets on Fire game	MusicGlove	Game score	BBT	Positive

FMA Fugl-Meyer Assessment, *ARAT* Action Research Arm Test, *MAS* Modified Ashworth Scale, *BBT* Box and Blocks Test, *TEMPA* Upper Extremity Performance Test for the Elderly, *WMFT* Wolf Motor Function Test, *MMSE* Mini Mental State Examination, *MAL* Motor Activity Log, *CAHAI* Chedoke Arm and Hand Activity Inventory, *MAL* Motor Activity Log

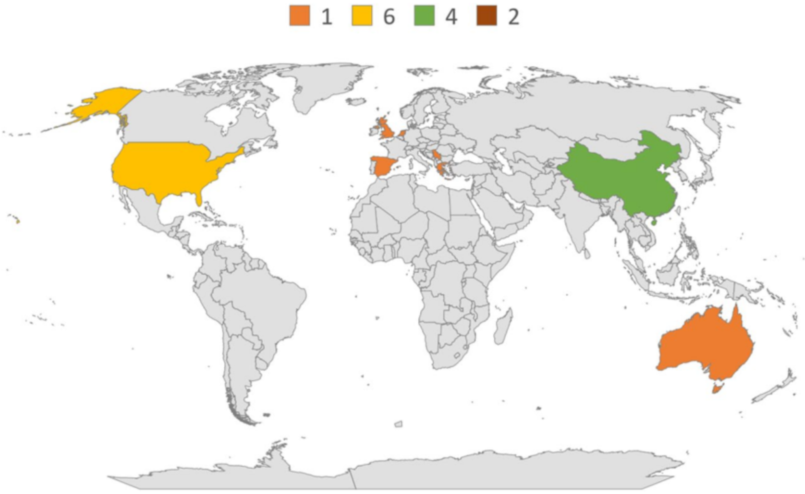


Fig. 2 Frequency distribution of studies by country. Alt Text: A world map with coloured countries, the darker the country is coloured, higher the frequency of studies. The majority of the studies came from the USA

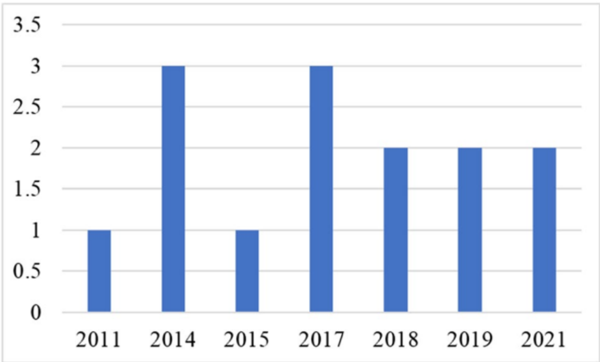


Fig. 3 Frequency distribution of studies by year. Alt Text: A bar chart showing the frequency of studies per year. 2017 and 2014 had the most studies, with three each

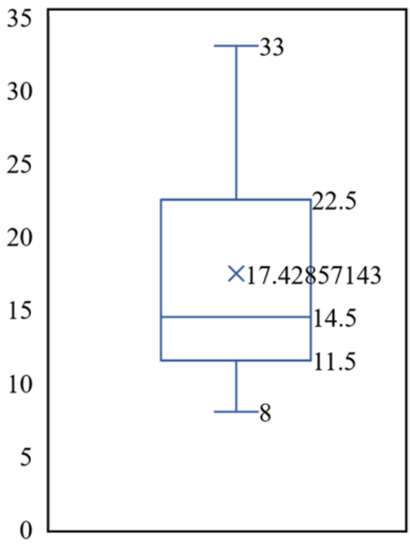


Fig. 4 Box plot of number of participants in the studies. Alt Text: A box-plot chart with the number of participants in the studies. The maximum number was 33, the minimum was 8, and the median was 17.4

Serbia and the UK had a single study represented. The distribution by country can be found in Fig. 2. It can be seen that there is a concentration of studies in few countries, with only 10 countries being represented, even though they are from four different continents.

Studies are also well distributed between journal and conferences, with 50% and 50% (7 studies for each method of publication). Regarding the year published, on average, the articles were published in 2016, with a standard deviation of approximately 3 years, which shows a relatively even distribution per year (Fig. 3).

Concerning the number of participants in the studies, a total of 244 individuals were involved. The study by Serradilla et al. [36] had the highest number, with 33 subjects, and the study by Astrakas et al. [25] had the lowest number, with eight participants. The average number was 17.42 (Fig. 4), which shows that the data from all studies must be applied to a larger population to be statistically confirmed for a relevant group of people.

4.3 Clinical scales

Nine different clinical scales were used to correlate data with. The most frequently used were the Fugl-Meyer Assessment (FMA) [21], appearing in eight studies, followed by the Box and Blocks Test (BBT) [23] and Wolf Motor Function Test (WMFT) [38], with four uses, each, and the Action Research Arm Test (ARAT) [39], with three uses. This can be better visualized graphically in Fig. 5.

Regarding the most popular clinical scales found, the FMA is an assessment index, specifically designed for post-stroke hemiplegic patients, that comprises five domains, with 155 items total. They are: (i) motor

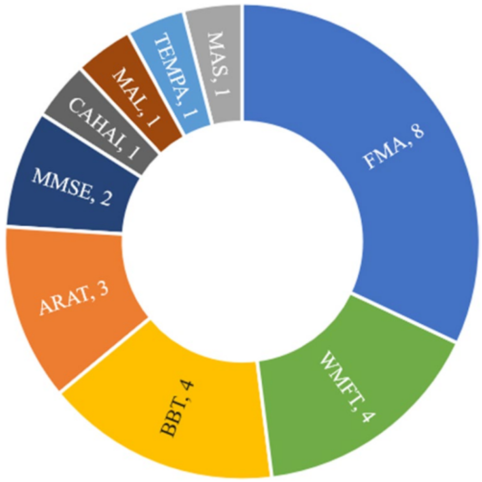


Fig. 5 Frequency distribution of clinical scales used in the studies. Alt Text: A pie chart showing the clinical scales most used in the studies. The most frequent were FMA (8 times), WMFT and BBT (4 times)

functioning; (ii) sensory functioning; (iii) balance; (iv) joint range of motion; and (v) joint pain. This scale requires several different tools, and approximately 30–35 min to be administered. It results in classifications for impairment severity, out of 100 points, being 100 the normal movement [21]. It has a significant amount of evidence proving its validity and good reliability, based on several studies in the literature, carried out over the years since the proposal of the FMA, in 1975 [40, 41].

Moreover, the ARAT is an assessment scale comprised of 19 items, proposed in 1981. Its items are categorized into four domains: (i) grasp; (ii) grip; (iii) pinch; and (iv) gross movement. The performance is rated on a 4-point scale, ranging from 0 to 3, being 3 the normal movement [39]. This test has also several studies supporting its validity. Furthermore, the ARAT has high concurrent validity with the FMA for upper limbs [42].

Another popular scale was the BBT. This test differs from the FMA and the ARAT in that it is not a questionnaire or an observational score, but a coordination exercise, which determines the score of the patient directly. The components to apply the BBT are: a wooden box, divided in two compartments, and 150 blocks. The individual is instructed to move as many blocks as possible, one at a time, from one compartment to the other for a period of 60 s. The number of blocks carried over to the other compartment is the patient's score of the BBT [23]. This scale is also well-known for upper-limb assessment post-stroke. Since its proposal, in 1985, it has been validated by several studies. For instance, Desrosiers et al. [43] showed that the BBT score has an adequate convergent validity with the FMA and excellent convergent validity with the ARAT.

Another scale that stood out was the WMFT, which consists of 21 items, separated in three parts: (i) time; (ii) functional ability; and (iii) strength. Tasks, such as "lift a can", are given, which must be completed as quickly as possible, with a time limit of 120 s. Then, a questionnaire must be filled, generating a 6-point score, from 0 to 5, with 5 being the normal movement. The WMFT was initially proposed in 2001, and showed agreement with the FMA score [38].

As found in the present review, other studies in literature also show that the FMA, ARAT, BBT and WMFT scales are popular for measuring post-stroke rehabilitation progress. For instance, a systematic literature review by Santisteban et al. [44] found a similar distribution, with the FMA being the most commonly used scale (36% of the studies), followed by the WMFT (19%). The ARAT (18%) was positioned as the 5th most popular, and the BBT (12%), as the 9th. This review also showed that only 15 measures were used in more than 5% of studies, which is similar to the distribution found here, that shows a tendency of only a few top scales to be used repeatedly.

In addition to the FMA, ARAT, BBT and WMFT, five scales were found, less than two times each, which are: MMSE (Mini-Mental State Examination) [45], CAHAI (Chedoke Arm and Hand Activity Inventory) [46], MAL (Motor Activity Log) [47], TEMPA (Upper Extremity Performance Test for the Elderly) [48] and MAS (Modified Ashworth Scale) [49]. Among these, only the MAS (listed as "Ashworth") and MAL were found as "popular" in [44], that is, were used in more than 5% of the studies.

4.4 Interaction methods

Motion tracking devices are used as input for rehabilitation exercises in a VR game. Many technologies and devices have been tested in recent years in post-stroke rehabilitation games [50]. These tracking devices can be classified into: mechanical, magnetic, vision (depth-based, marker-based, markerless), inertial and ultrasonic.

Among the interaction methods found in the chosen studies, the most popular were the following: (i) Microsoft Kinect [51] and Leap Motion [52], depth cameras systems capable of skeleton tracking, used five times [3, 5, 6, 14, 20]; (ii) inertial sensor or Wii remote controller [52], motion sensing devices, such as accelerometers, used three times [3, 10, 21]; (iii) exoskeletons or other mechanical robotic devices, used three times [2, 7, 41]; (iv) touch screen, used three times [11–13]; (v) other devices: the "MusicGlove", a custom glove with electrical contacts on the fingertips, used two times (by the same author) [19, 53]; and (vi) magnetic motion sensors, used once [36]. These results are in line with the current literature, which cites Kinect, Leap Motion and Wii among the most popular motion trackers for post-stroke rehabilitation [54, 55].

4.5 User's data parameters from the games

This review sought studies that assessed the validity of user data parameters obtained from game software by correlating them with a clinical scale. Therefore, any data acquired through the game software during the user's engagement with it, regardless of the type of clinical study, was considered in this review if it was also obtained in conjunction with a relevant clinical physiotherapy scale for the same patient. Only the scales that had some form of upper limb test in them were considered. In this way, it can be compared, for the same population, if the data obtained from the software and the score from the clinical scale are positively correlated. As mentioned previously, this correlation calculation was an inclusion criterion for the study. If the calculation was not made, the study was excluded from this review.

Therefore, in the studies found, among the virtual data captured while the user plays, some different categories could be noticed and separated. They were, in order of most used: game score, used in eight studies; kinematic parameters, such as the speed and smoothness of movement, used in six studies; and time variables, such as shortest time to finish a level, used in three studies (Table 3).

4.6 Correlation with clinical scales

When reporting the correlation between the clinical scale(s) used and the measurements obtained while using the software, 12 of the 14 studies reported positive or strongly positive correlations, which represents 85.7% of all studies. Only the study by Cidota et al. [32] showed a negative correlation and the one by Lee et al. [26] showed mixed correlations, that is, positive for some parameters and negative for others (Fig. 6).

4.6.1 Game score

In the work of [25], it was tested a robot-assisted therapy, consisting of a labyrinth serious game, in combination with a robotic device, MR_CHIROD, which is a newly re-designed robotic hand rehabilitation device, engineered to provide adjustable levels of force for handgrip exercising. In this study, the following serious game elements were collected: number of collisions, rewards and total score. Also, the clinical scales utilized were the FMA, ARAT, BBT and MAS. They observed several strong correlations between clinical motor scale scores with game metrics. Among those, the strongest correlations

Table 3 Categories of the assessment parameters of the games

Game score	[25, 27, 29–31, 34, 35, 37]
Kinematic parameters	[17, 26, 28, 32, 33, 36]
Time variables	[27, 29, 32]

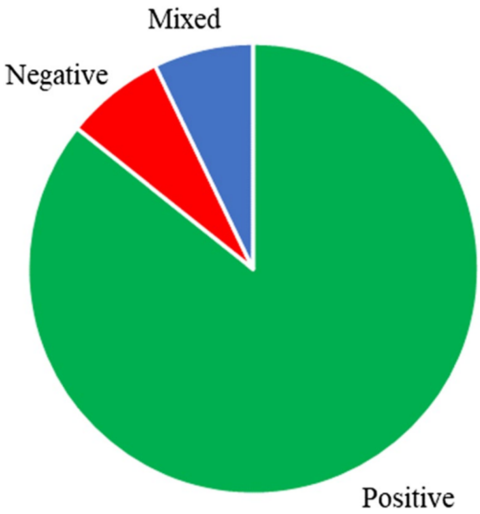


Fig. 6 Correlation results of the in-game parameters with clinical scales. Alt Text: A pie chart showing the correlation results of game parameters with clinical scales. Most of the results were positive

found were the number of collisions with the BBT ($\rho = -0.670$, $p < 0.001$), the number of collisions with the FMA ($\rho = -0.642$, $p < 0.001$), followed by the score with the BBT ($\rho = 0.636$, $p < 0.001$), where ρ is the Spearman correlation coefficient and p is the p-value.

Another study to find that the game score was positively correlated with the FMA was [30]. However, they used a different form of interaction than [25]. This time, an inertial sensor integrated in a motion tracking device was used. They found a significant positive correlation between game score and the FMA in one of the groups ($r = 0.86$, $p = 0.03$), where r is the Pearson correlation coefficient and p is the p-value.

Moreover, the study by [31] also tested the FMA to game score correlation. In their work, however, another form of interaction was utilized, which is the Wiimote, even though the Wiimote can be considered in the inertial sensor category. They demonstrated that clinical assessments and game score demonstrated improved motor function for all patients at post-therapy ($p < 0.01$), where p is the p-value, so that was considered a positive correlation for this review.

In another study [35], the Leap Motion controller was used to interact with the serious game. Again, the game score and the FMA were compared. The game score showed high correlation with both FMA ($r = 0.72$, $p < 0.05$), and BBT ($r = 0.86$, $p < 0.01$) scores, where r is the Pearson correlation coefficient and p is the p-value. Moreover, they suggested that the game score could be a good indicator of the hand function due to these results.

Furthermore, some studies compared the game score with other clinical scales. The studies [27] and [29], both by the same author, used the MMSE. The form of interaction was a touch screen. They calculated the normalized root mean square error between the actual and estimated MMSE scores, obtaining a result of 5.3%, along with machine learning tests, which was considered as a good probability that their game could predict a MMSE score accurately.

Moreover, the studies [34] and [37], both by the same author, used a proprietary electronic glove device to interact with the game. They compared the game and the BBT scores, finding a significant linear relationship between them ($p < 0.001$, $r = 0.89$), where r is the Pearson correlation coefficient and p is the p-value, which is also a strong correlation.

4.6.2 Kinematic parameters

Kinematic parameters are measurements obtained using electronic devices capable of capturing specific body movements. For example, the Microsoft Kinect can track body coordinates, including joint speeds and angles. Other devices, such as inertial sensors, exoskeletons, and artificial intelligence-aided cameras, also enable the capture of such kinematic data.

The study [26], for instance, measured some kinematic parameters such as movement offset, speed and efficiency, with Microsoft Kinect. They compared the kinematic parameters with FMA, TEMPA and WMFT scores, using the Spearman correlation. As a result, several of the motor indicators were highly correlated with the evaluation scales. For example, the correlation between movement speed and the FMA was positive ($\rho = 0.807$, $p < 0.01$).

The study by [33] specifically sought to identify the motion kinetics extracted from the VR based rehabilitation that were significantly correlated with the functional improvement. They also used Kinect for capturing kinematic data, such as mean speed, velocity and smoothness. They also tested the correlation against FMA and WMFT, therefore being similar to [26]. They stated that the correlation was positive, but did not inform the correlation coefficients.

Another study that used kinematic parameters and Kinect was [32]. Unlike the majority of the studies found, this used an Augmented Reality game, instead of a VR one. As kinematic parameters, they used the ARAT test movements to capture joint velocities, angles and trajectories. This was the only study in this review to find no significant correlations between the movements captured in real world (ARAT) and movements in the virtual world. They did not inform the specific correlation coefficients.

Moreover, the study by [17] tested kinematic variables against FMA, ARAT and WMFT, which were popular in this category. However, instead of Kinect, they used a low-cost robotic system to capture movement. A correlation analysis between the clinical scales and the different games' scores, kinematic and kinetic measures extracted from robot was done. The results from this test were mostly statistically significantly correlated with the clinical scales, for example, the range of movement measured with the FMA and with their system had a correlation of 0.68, $p < 0.01$.

However, the study [36] was different. Firstly, they were the only one to use the CAHAI test. Also, they used magnetic motion tracking, which wasn't popular among the studies. Moreover, they made a complex form of analysis, which involved 320 kinematic variables, to characterize how well the movement was performed, such as speed, fluency and synchrony. They could demonstrate cross-sectional validity with great correlation coefficient ($r = 0.998$, $p < 0.001$); and longitudinal validity within the subjects, with good correlation ($r = 0.54$, $p < 0.001$).

Moreover, the study by [28] used a sensor fusion method, based on Kinect and a inertial sensor to capture motion information. They used variables not found in other studies, such as elevation angle and azimuth angle of joints relative to other points. Also, they did specifically the movements requested by the FMA, while recording virtually with the inertial sensor. They found these data had a good correlation with the real result from the physician ($r = 0.902$, $p < 0.001$), where r is the Pearson correlation coefficient and p is the p-value.

4.6.3 Time variables

Besides the game scores and kinematic parameters, variables that were time-related were also found. This include time to finish a phase of the game, for instance.

The studies by [27] and [29], mentioned previously, also used time variables, mixed in with the game score, such as shortest time, longest time and mean time to clear the stages. They didn't specify the time variables correlations, but only provided information about the total correlation (time + game score variables), which was positive.

The study by [32] used the Spearman's rho (ρ) to investigate if characteristics of movements in Augmented Reality were related to movements in the real world, captured with the ARAT. They stated that no significant correlations were observed between them, not informing the specific values of ρ.

5 Discussion

Due to their various advantages over traditional therapies, VR-based systems have the potential to be used as physical intervention tools. These benefits include their affordability, mobility, and interoperability with other systems, in addition to the capacity to give patients immediate feedback and the opportunity to employ games to motivate participants [56]. Additionally, the use of VR in home-based therapies may be successful in stroke recovery [57].

The data gathered in this review made possible to notice that multiple combinations of interaction devices, data parameters and assessment scales were used. Also, it was found an expressive predominance (85.7%) of favorable correlations between the virtual data captured during interaction with the VR system and the real data obtained from traditionally used clinical measurements. However, the combinations between devices and scales were so varied that it is not possible to affirm a specific pair of features that performed consistently. Nevertheless, some combinations found repeatedly were the game score parameters and FMA scores, found in four studies, and Kinect and FMA parameters, found in three. It was also possible to tell that the game scores included several different forms of motion capture, while the kinematic parameters were frequently captured with Kinect, in four from the six studies from this category [26, 28, 32, 33].

Among all the studies, only one had found no correlations between clinical measurements (ARAT) and the game utilized [32]. They used Kinect and the ARAT test, both popular options among the studies. The only difference between this and the other studies, was that this used an Augmented Reality game, instead of a VR one. However, this cannot be used to affirm that this always brings a negative correlation, because there is not enough evidence or an obvious relationship between them.

Therefore, it would be interesting to analyze which specific parameters always correlate with specific scales, in order to find rules and patterns to be applied, preventing errors. Also, it was verified the need for more participants in the studies, which is necessary to make a statistically relevant statement to be applied to a larger population.

6 Conclusion

This scoping review's objective was to assess VR games and simulations for post-stroke rehabilitation that recorded data from the users while they were playing and correlated the acquired information with traditional physiotherapy scales. The data parameters could be classified into three categories: game score (eight studies), kinematic parameters (six studies) and time variables (three studies). Moreover, the clinical scales used the most were the FMA (eight studies), the WMFT and BBT (four studies each). The great majority of the retained studies (12/14) found positive correlation between these data and the well-known scales. Combinations of game variables and physiotherapy scales commonly used were the game scores and FMA and kinematic parameters from Kinect and FMA.

In this way, we can suggest that some game parameters, in particular, game score, kinematic parameters and time variables can be linked to the design of games and simulators in VR for post-stroke rehabilitation and be treated with greater importance, since all of them showed potential clinical validity, in the different setups and combinations.

However, researches using a larger sample size and a variety of software, scales, and parameters are needed to establish that parameters coming from assessment in VR have the same validity as the actual testing scales. As an initial effort, this study can point to a promising direction.

Finally, according to Pietrzak et al. [58], the need to incorporate video games based on VR systems in stroke rehabilitation should be emphasized, in addition to facilitating the use of the service for therapy centers and home treatments. Thus, their use could become common practice, in addition to potentially assisting in the clinical assessment of motor performance, virtually and in real time, improving practice and saving time and resources for patients and therapists.

Acknowledgements The authors would like to acknowledge the support of CAPES (Coordination for the Improvement of Higher Education Personnel), through scholarships and research support.

Authors' contributions JT was involved in drafting the manuscript and revising it critically for important intellectual content. ELMN and YM supervised, revised and gave the final approval of the manuscript. All authors were fully involved in the study and preparation of the manuscript. All authors read and approved the final manuscript.

Funding This study was financed in part by the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) through scholarship.

Data availability All data generated or analyzed during this study are included in this published article.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Disclosure of interest The authors declare that they have no competing interests.

References

1. Tsao CW, Aday AW, Almarzooq ZI et al (2022) Heart Disease and Stroke Statistics-2022 Update: A Report From the American Heart Association. Circulation 145:e153–e639
2. Lohse KR, Hilderman CGE, Cheung KL et al (2014) Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. PLoS One 9:e93318
3. Kiper P, Agostini M, Luque-Moreno C et al (2014) Reinforced feedback in virtual environment for rehabilitation of upper extremity dysfunction after stroke: preliminary data from a randomized controlled trial. Biomed Res Int 2014:752128
4. Murie-Fernández M, Irimia P, Martínez-Vila E et al (2010) Neurorehabilitation after stroke. Neurologia 25:189–196

5. Broeren J, Claesson L, Goude D et al (2008) Virtual rehabilitation in an activity centre for community-dwelling persons with stroke. The possibilities of 3-dimensional computer games. Cerebrovasc Dis 26:289–296
6. Merchant Z, Goetz ET, Cifuentes L et al (2014) Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. Comput Educ [Internet] 70:29–40. Available from: https://www.sciencedirect.com/science/article/pii/S0360131513002108
7. Norouzi-Gheidari N, Hernandez A, Archambault PS et al (2020) Feasibility, Safety and Efficacy of a Virtual Reality Exergame System to Supplement Upper Extremity Rehabilitation Post-Stroke: A Pilot Randomized Clinical Trial and Proof of Principle. Int J Environ Res Public Health 17
8. Pourmand A, Davis S, Marchak A et al (2018) Virtual reality as a clinical tool for pain management. Curr Pain Headache Rep 22:1–6
9. Yeh S-C, Chen Y-C, Tsai C-F et al (2012) An innovative virtual reality system for mild cognitive impairment: Diagnosis and evaluation. 2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences, Langkawi, Malaysia, pp 23–27. https://doi.org/10.1109/IECBES.2012.6498023
10. Barteit S, Lanfermann L, Bärnighausen T et al (2021) Augmented, mixed, and virtual reality-based head-mounted devices for medical education: systematic review. JMIR Serious Games 9:e29080
11. Pourmand A, Davis S, Lee D et al (2017) Emerging Utility of Virtual Reality as a Multidisciplinary Tool in Clinical Medicine. Games Health J 6:263–270
12. Park Y-H, Lee C-H, Lee B-H (2013) Clinical usefulness of the virtual reality-based postural control training on the gait ability in patients with stroke. J Exerc Rehabil 9:489–494
13. Weber LM, Nilsen DM, Gillen G et al (2019) Immersive virtual reality mirror therapy for upper limb recovery following stroke: A pilot study. Am J Phys Med Rehabil 98:783
14. Ahmad MA, Singh DKA, Mohd Nordin NA et al (2019) Virtual reality games as an adjunct in improving upper limb function and general health among stroke survivors. Int J Environ Res Public Health 16(24):5144. https://doi.org/10.3390/ijerph16245144
15. Kim W-S, Cho S, Baek D et al (2016) Upper Extremity Functional Evaluation by Fugl-Meyer Assessment Scoring Using Depth-Sensing Camera in Hemiplegic Stroke Patients. PLoS One 11:e0158640
16. Adams RJ, Ellington AL, Armstead K et al (2019) Upper Extremity Function Assessment Using a Glove Orthosis and Virtual Reality System. OTJR (Thorofare N J) 39:81–89
17. Rodriguez-de-Pablo C, Balasubramanian S, Savic A et al (2015) Validating ArmAssist Assessment as outcome measure in upper-limb post-stroke telerehabilitation. Annu Int Conf IEEE Eng Med Biol Soc 2015:4623–4626. https://doi.org/10.1109/EMBC.2015.7319424
18. Henderson A, Korner-Bitensky N, Levin M (2007) Virtual reality in stroke rehabilitation: a systematic review of its effectiveness for upper limb motor recovery. Top Stroke Rehabil 14:52–61
19. Kwon J-S, Park M-J, Yoon I-J et al (2012) Effects of virtual reality on upper extremity function and activities of daily living performance in acute stroke: a double-blind randomized clinical trial. NeuroRehabilitation 31:379–385
20. Laver KE, Lange B, George S et al (2017) Virtual reality for stroke rehabilitation. Cochrane Database Syst Rev 9:CD008349. https://doi.org/10.1002/14651858.CD008349.pub2
21. Fugl-Meyer AR, Jääskö L, Leyman I et al (1975) A method for evaluation of physical performance. Scand J Rehabil Med 7:13–31
22. Wolf SL, Thompson PA, Morris DM et al (2005) The EXCITE Trial: Attributes of the Wolf Motor Function Test in Patients with Subacute Stroke. Neurorehabil Neural Repair [Internet] 19:194–205. Available from: https://doi.org/10.1177/1545968305276663
23. Mathiowetz V, Volland G, Kashman N et al (1985) Adult Norms for the Box and Block Test of Manual Dexterity. Am J Occup Ther 39:386–391
24. Page MJ, McKenzie JE, Bossuyt PM et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 10:89. https://doi.org/10.1186/s13643-021-01626-4
25. Astrakas LG, De Novi G, Ottensmeyer MP et al (2021) Improving motor function after chronic stroke by interactive gaming with a redesigned MR-compatible hand training device. Exp Ther Med 21:245
26. Lee S-H, Cui J, Liu L et al (2021) An Evidence-Based Intelligent Method for Upper-Limb Motor Assessment via a VR Training System on Stroke Rehabilitation. IEEE Access 9:65871–65881
27. Jung H-T, Daneault J-F, Lee H et al (2019) Remote Assessment of Cognitive Impairment Level Based on Serious Mobile Game Performance: An Initial Proof of Concept. IEEE J Biomed Heal Inform 23:1269–1277
28. Bai J, Song A (2019) Development of a Novel Home Based Multi-Scene Upper Limb Rehabilitation Training and Evaluation System for Post-Stroke Patients. IEEE Access 7:9667–9677
29. Jung H-T, Lee H, Kim K et al (2018) Estimating Mini Mental State Examination Scores using Game-Specific Performance Values: A Preliminary Study. Conf Proc Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf 2018:1518–1521
30. Lin B, Chen J, Hsu H (2018) Novel Upper-Limb Rehabilitation System Based on Attention Technology for Post-Stroke Patients: A Preliminary Study. IEEE Access 6:2720–2731
31. Hesam-Shariati N, Trinh T, Thompson-Butel AG et al (2017) A Longitudinal Electromyography Study of Complex Movements in Poststroke Therapy. 2: Changes in Coordinated Muscle Activation. Front Neurol 8:277
32. Cidota MA, Bank PJM, Ouwehand PW et al (2017) Assessing Upper Extremity Motor Dysfunction Using an Augmented Reality Game. 2017 IEEE Int Symp Mix Augment Real. p. 144–154
33. Chen C, Lee S, Wang W et al (2017) The changes of improvement-related motor kinetics after virtual reality based rehabilitation. 2017 Int Conf Appl Syst Innov. p. 683–685
34. Friedman N, Chan V, Reinkensmeyer AN et al (2014) Retraining and assessing hand movement after stroke using the MusicGlove: comparison with conventional hand therapy and isometric grip training. J Neuroeng Rehabil 11:76
35. Khademi M, Mousavi Hondori H, McKenzie A et al (2014) Free-Hand Interaction with Leap Motion Controller for Stroke Rehabilitation. CHI '14 Ext Abstr Hum Factors Comput Syst [Internet]. Association for Computing Machinery, New York, NY, USA. p. 1663–1668. Available from: https://doi.org/10.1145/2559206.2581203
36. Serradilla J, Shi JQ, Cheng Y et al (2014) Automatic assessment of upper limb function during play of the action video game, circus challenge: validity and sensitivity to change. 2014 IEEE 3nd Int Conf Serious Games Appl Heal. 1–7
37. Friedman N, Chan V, Zondervan D et al (2011) MusicGlove: motivating and quantifying hand movement rehabilitation by using functional grips to play music. Conf Proc Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf 2011:2359–2363
38. Wolf SL, Catlin PA, Ellis M et al (2001) Assessing Wolf motor function test as outcome measure for research in patients after stroke. Stroke 32:1635–1639
39. Yozbatiran N, Der-Yeghiaian L, Cramer SC (2007) A standardized approach to performing the action research arm test. Neurorehabil Neural Repair 22(1):78–90. https://doi.org/10.1177/1545968307305353
40. Gladstone DJ, Danells CJ, Black SE (2002) The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil Neural Repair 16:232–240

41. Duncan PW, Propst M, Nelson SG (1983) Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. Phys Ther 63:1606–1610
42. McDonnell M (2008) Action research arm test. Aust J Physiother 54:220
43. Desrosiers J, Bravo G, Hébert R et al (1994) Validation of the Box and Block Test as a measure of dexterity of elderly people: reliability, validity, and norms studies. Arch Phys Med Rehabil 75:751–755
44. Santisteban L, Térémetz M, Bleton JP et al (2016) Upper Limb Outcome Measures Used in Stroke Rehabilitation Studies: A Systematic Literature Review. PLoS One 11:e0154792
45. Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189–198
46. Barreca S, Gowland CK, Stratford P et al (2004) Development of the Chedoke Arm and Hand Activity Inventory: theoretical constructs, item generation, and selection. Top Stroke Rehabil 11:31–42
47. Taub E, Miller NE, Novack TA et al (1993) Technique to improve chronic motor deficit after stroke. Arch Phys Med Rehabil 74:347–354
48. Desrosiers J, Hébert R, Bravo G et al (1995) Upper extremity performance test for the elderly (TEMPA): normative data and correlates with sensorimotor parameters. Test d'Evaluation des Membres Supérieurs de Personnes Agées. Arch Phys Med Rehabil 76:1125–1129
49. Bohannon RW, Smith MB (1987) Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther 67:206–207
50. Wang Q, Markopoulos P, Yu B et al (2017) Interactive wearable systems for upper body rehabilitation: a systematic review. J Neuroeng Rehabil 14:20
51. Microsoft (2021) Azure Kinect DK [Internet]. [cited 2021 Jul 2]. Available from: https://www.microsoft.com/en-us/d/azure-kinect-dk/8pp5vxmd9nhq?activetab=pivot:overviewtab
52. Ultraleap (2023) Leap Motion Controller [Internet]. [cited 2023 Mar 13]. Available from: https://www.ultraleap.com/product/leap-motion-controller/
53. Adamovich SV, Fluet GG, Tunik E et al (2009) Sensorimotor training in virtual reality: a review. NeuroRehabilitation 25:29–44
54. Alex M, Chen C, Wunsche BC (2017) A review of sensor devices in stroke rehabilitation. 2017 International Conference on Image and Vision Computing New Zealand (IVCNZ), Christchurch, New Zealand, pp 1–6. https://doi.org/10.1109/IVCNZ.2017.8402480
55. Thomson K, Pollock A, Bugge C et al (2016) Commercial gaming devices for stroke upper limb rehabilitation: a survey of current practice. Disabil Rehabil Assist Technol 11:454–461
56. Ruiz-González L, Lucena-Antón D, Salazar A et al (2019) Physical therapy in Down syndrome: systematic review and meta-analysis. J Intellect Disabil Res 63:1041–1067
57. Lin J, Kelleher CL, Engsberg JR (2013) Developing Home-Based Virtual Reality Therapy Interventions. Games Health J 2:34–38
58. Pietrzak E, Pullman S, McGuire A (2014) Using Virtual Reality and Videogames for Traumatic Brain Injury Rehabilitation: A Structured Literature Review. Games Health J 3:202–214

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Julia Tannus received her Master of Science degree in 2019, in computer engineering. She is a PhD student and substitute professor at Federal University of Uberlandia—Brazil.

Eduardo L. M. Naves received his Master of Science degree in 2001 and Doctor of Science in 2006, both in the area of biomedical engineering. He is a professor and researcher (Federal University of Uberlandia—Brazil).

Yann Morere received his Master of Science degree in 1996 and Doctor of Science in 2001, both in the area of industrial and human automatic system. He is a professor and researcher (Université de Lorraine).

2.1.2 Key Findings and Implications

The reproduced article provides a comprehensive overview of the current landscape of VR-based and game-based motor assessment in post-stroke rehabilitation, highlighting both the potential and the limitations of existing approaches. Its findings justify the research problem addressed in this thesis.

In summary, the review demonstrates that game-derived parameters can achieve clinically meaningful correlations with established assessment scales, suggesting that rehabilitation games have potential to serve not only as therapeutic tools but also as instruments for objective functional evaluation. Among the different types of parameters analyzed, kinematic features and game performance scores emerged as particularly relevant indicators of motor function.

However, the study also reveals that most existing solutions depend heavily on external sensors, complex hardware setups, or non-transparent computational models. These factors limit their applicability in real-world clinical settings, especially in scenarios involving remote monitoring or low-resource environments. Furthermore, the lack of standardization across studies and the predominance of small-scale experiments highlight the need for more scalable, accessible, and interpretable solutions.

Within the context of this thesis, these findings directly informed the design choices of the proposed system. They motivated the transition from hardware-dependent approaches to a vision-based, sensor-free framework, as well as the adoption of interpretable kinematic features and lightweight modeling techniques. By addressing the limitations identified in the literature, the subsequent phases of this research aim to advance the field toward more practical and clinically applicable solutions.

2.2 Pilot test

2.2.1 Context and Summary

This section presents the full reproduction of the article entitled *Low-Cost Vision-Based 3-D Elbow Tracking for Post-Stroke Rehabilitation: Development and Pilot Evaluation of a Serious Game*, published in IEEE Transactions on Neural Systems and Rehabilitation Engineering (TANNUS *et al.*, 2025). This article is distributed under a Creative Commons Attribution 4.0 International License, which permits reproduction, distribution, and adaptation, provided appropriate credit is given to the original authors.

This work corresponds to Phase 2 of the research objectives of this doctoral thesis and represents the transition from theoretical investigation to practical system development. While the previous study established the clinical relevance of game-derived metrics for motor assessment, this article focuses on the design, implementation, and validation of a low-cost technological solution capable of capturing clinically meaningful movement data during rehabilitation exercises.

The main contribution of this study is the development of a vision-based exergame system for real-time 3D elbow angle estimation, using only a standard RGB camera and simple physical markers. Unlike conventional approaches that rely on depth sensors, wearable devices, or robotic systems, the proposed method employs color segmentation and geometric reconstruction techniques to estimate joint angles without specialized hardware.

The system architecture integrates four main modules: camera calibration, color-based marker detection, an interactive exergame environment, and a motion report module. Together, these components enable real-time tracking, adaptive gameplay, and automated performance feedback. The exergame was designed to promote repetitive elbow flexion and extension movements within an engaging and task-oriented virtual environment.

To evaluate the proposed solution, a comprehensive experimental protocol was conducted, including technical validation and usability testing. The results demonstrated that the system achieves high angular accuracy (error below 5°) for clinically relevant movement ranges up to approximately 110°, when compared to a goniometer reference. Additional analyses showed robustness across different calibration conditions, marker sizes, and camera distances, as well as predictable degradation under high-speed motion or peripheral image regions.

Furthermore, a comparative evaluation against the AI-based MediaPipe framework revealed that the proposed method provides more stable and reliable estimations in occlusion scenarios, particularly when the arm is positioned close to the torso. This highlights the advantage of the geometric approach over purely AI-based pose estimation in specific rehabilitation contexts.

A usability study conducted with eight post-stroke participants resulted in a mean System Usability Scale (SUS) score of 92.5, indicating excellent user acceptance and suggesting the feasibility of the system for real-world and home-based rehabilitation applications.

In conclusion, the central idea of this work is that it is possible to accurately measure joint movement in three dimensions using only a simple camera and low-cost markers, enabling accessible and scalable rehabilitation solutions without the need for expensive equipment.

Low-Cost Vision-Based 3-D Elbow Tracking for Post-Stroke Rehabilitation: Development and Pilot Evaluation of a Serious Game

Julia Tannus, Camille Alves, Caroline Valentini, Yann Morere, Guy Bourhis, Pierre Pino, and Eduardo Naves

Abstract*— Stroke is a leading contributor to long-term disability worldwide, and rehabilitation often relies on costly devices, limited infrastructure, or labor-intensive protocols. While virtual reality-based exergames have gained popularity for promoting patient engagement, most rely on proprietary sensors or wearable electronics, limiting accessibility and clinical adaptability. This study presents the design, implementation, and pilot evaluation of a custom exergame that estimates the 3D elbow angle using a single RGB camera and two colored spheres as markers, eliminating the need for specialized hardware. The proposed system performs camera calibration, color segmentation, geometric 3D reconstruction, and real-time elbow angle estimation using low-cost equipment. Extensive technical tests revealed robust performance, with angular errors below 5° for joint amplitudes under 110°, and consistent accuracy across different lighting conditions, marker sizes, and distances. Additional tests showed that excessive sphere velocity (>20 cm/s) or proximity to image corners increased error due to motion blur and lens distortion, respectively. The system outperformed the AI-based MediaPipe framework in occluded-arm scenarios. Regression analysis showed strong correlation (r = 0.70) between movement velocity and angular error. Usability testing with eight post-stroke participants yielded a mean SUS score of 92.5/100. The proposed solution is a promising alternative for home-based, sensor-free rehabilitation, supporting personalized exercise routines and remote progress monitoring.

***Index Terms*— 3D tracking, computer vision, elbow joint, post-stroke rehabilitation, serious games, telerehabilitation, user-centered design.**

Received 30 April 2025; revised 25 June 2025; accepted 15 July 2025. Date of publication 21 July 2025; date of current version 31 July 2025. This work was supported in part by the National Council for Scientific and Technological Development (CNPq), Brazil, under Grant 440439/2024-8, Grant 400780/2024-0, and Grant 445256/2024-9; and in part by the Coordination for the Improvement of Higher Education Personnel (CAPES) and the Financial Agency for Studies and Projects (FINEP) under Grant 01.24.0116.00. The work of Julia Tannus was supported by CAPES, Brazil, under Grant 88887.663238/2022-00. The work of Camille Alves was supported by CNPq, Brazil, under Grant 153088/2024-9. The work of Eduardo Naves was supported by CNPq, Brazil, under Grant 303614/2023-4. *(Corresponding author: Julia Tannus.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Federal University of Uberlandia under Protocol No. 39232820.2.0000.5152.

Julia Tannus, Camille Alves, Caroline Valentini, and Eduardo Naves are with the Assistive Technologies Group, Federal University of Uberlandia, Uberlandia 2121, Brazil (e-mail: julia.tannus@ufu.br; eduardonaves@ufu.br).

Yann Morere, Guy Bourhis, and Pierre Pino are with the LCOMS Laboratory, University of Lorraine, 57000 Metz, France.

Digital Object Identifier 10.1109/TNSRE.2025.3591104

I. Introduction

STROKE is a major global health concern, recognized as the second leading cause of death and the third leading cause of disability [1]. One in four people will experience a stroke in their lifetime, regardless of sociodemographic background [2]. In 2020, a stroke occurred every 40 seconds in the United States [1]. Despite therapeutic efforts, 30% to 66% of stroke survivors experience persistent arm dysfunction after six months [3], [4], [5], and only 5% to 20% achieve complete functional recovery [6], [7]. This loss of upper limb function significantly compromises patients' autonomy and quality of life.

Although physical therapy is the cornerstone of motor rehabilitation, traditional programs are often time-consuming, resource-intensive, and limited in availability, especially in remote regions [6]. These programs can also be tedious and financially burdensome, requiring patient transportation to clinical sites. In contrast, video game-based rehabilitation has emerged as a motivational and accessible alternative [8], [9]. Virtual reality games have proven capable of engaging stroke patients in repetitive motor tasks that support neuroplasticity and motor recovery [10].

However, accessibility challenges persist. Many health professionals report barriers related to the cost and availability of therapeutic technology [9], often relying on commercial exergames like the Wii Remote [11] or Kinect [12], [13]—devices not originally designed for post-stroke patients. Additional solutions include wearable electronics, such as IMUs [14], [15], [16], LeapMotion [13], or robotic rehabilitation systems [17], [18], but these frequently involve high costs, maintenance, and limited adaptability. These limitations have prompted the development of custom low-cost alternatives [19], [20], [21].

Commercial games are also limited in therapeutic scope, often lacking personalized tasks, safety monitoring,

© 2025 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

or clinically validated performance metrics [9]. Consequently, stroke-specific solutions are needed. Accurate measurement of elbow range of motion is particularly relevant, as spasticity—a condition characterized by involuntary muscle contractions—affects 19% to 43% of post-stroke individuals [22]. Elbow angle tracking can aid in assessing motor recovery and tailoring rehabilitation [23], [24]. Also, it can help to detect compensatory movements such as torso adjustments, thus encouraging proper motor execution aligned with conventional rehabilitation protocols [25].

Depth-sensing cameras and wearable IMUs can track 3D joint motion effectively [24], [26], but reliance on proprietary hardware introduces concerns related to cost, obsolescence, and mobile device integration. These constraints have led researchers to explore vision-based systems that rely solely on camera input, especially those enhanced by Artificial Intelligence (AI), as scalable, low-cost alternatives for rehabilitation without the need for external sensors [27].

In recent years, arm tracking with AI has seen significant advancements [24], [28], [29], [30]. AI-powered systems are now capable of tasks such as hand tracking and real-time 3D pose estimation. Ready-to-use frameworks like MediaPipe [31] have emerged to facilitate implementation. However, capturing the elbow angle in 3D remains challenging for AI due to occlusions, depth ambiguity, and reliance on monocular image inputs, which often result in errors when estimating joint positions beyond a 2D plane [32], [33]. Additionally, frameworks such as MediaPipe do not provide real depth calculation but only relative depth from a reference point (0), limiting the accuracy of absolute 3D positioning. While relative 3D coordinates are sufficient for many interactive applications, clinical assessments and progress tracking require absolute spatial measurements (e.g., in centimeters or degrees) to evaluate range of motion, reach, and joint angles over time.

Furthermore, real-time performance requires considerable computational resources, primarily due to the need for deep neural networks that process high-dimensional feature spaces, perform complex pose estimation, and require GPU acceleration or specialized hardware for inference [24]. At times, traditional image processing, such as color segmentation, can solve a given problem without the need for Deep Learning calculations, thus making a software more computationally efficient [34].

In response to these challenges, this study presents an exergame that integrates a low-cost, vision-based 3D elbow tracking system. The proposed system was designed to be capable of running on the user's existing hardware, such as a smartphone, tablet, or gaming laptop, without requiring any additional sensors. The system uses two color-coded spheres, which can be produced using basic stationery materials, with an estimated cost of USD 5, and geometric depth projection to estimate elbow angle in real time. This flexibility and minimal material requirement qualify the solution as low-cost and suitable for scalable, home-based rehabilitation.

This custom-built game supports personalized rehabilitation exercises and includes a progress monitoring tool to assist therapists. Unlike previous solutions, it delivers absolute 3D joint measurements using only monocular RGB video, with no dependence on proprietary depth sensors or wearables. This work details the development and pilot evaluation of the system, aiming to validate its accuracy, feasibility on mobile devices, and potential for clinical use.

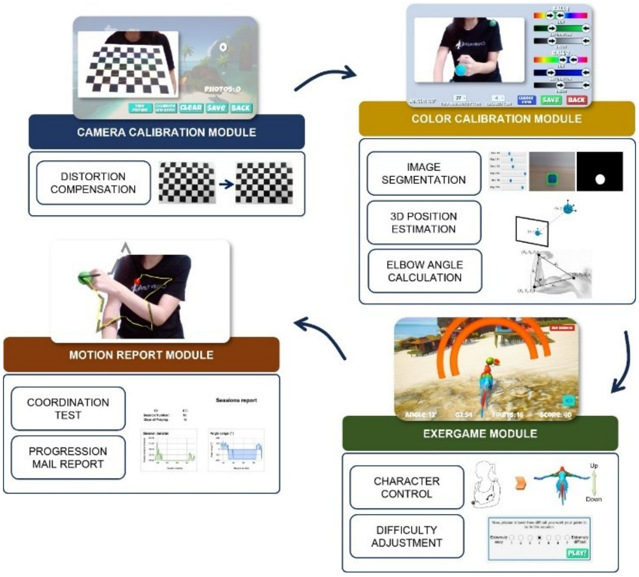


Fig. 1. Overview of the proposed system architecture. The pipeline includes four main modules: (1) *Camera Calibration Module*, which compensates lens distortion through checkerboard pattern analysis; (2) *Color Calibration Module*, where the system detects colored spheres, estimates their 3D position, and computes the elbow angle in real time; (3) *Exergame Module*, where the estimated angle controls the virtual character and gameplay difficulty is adjusted dynamically; and (4) *Motion Report Module*, which provides feedback reports including performance metrics and progress tracking.

II. Materials and Methods

This section will cover the materials and methods used for developing and testing the software.

A. System Overview and Architecture

The proposed system architecture follows the diagram shown in Fig. 1. The system comprises four interconnected modules: Camera Calibration, Color Calibration, Exergame, and Motion Report. The Camera Calibration Module ensures accurate motion capture by compensating for lens distortion using a checkerboard pattern. The Color Calibration Module processes the captured images through segmentation, estimates 3D positions, and calculates elbow angles based on detected markers. This data is fed into the Exergame Module, where patients control an on-screen character through their movements, with adjustable difficulty levels for personalized therapy. Finally, the Motion Report Module analyzes movement performance, providing coordination tests and progress reports, which can be sent via email. Together, these modules enable precise motion tracking and interactive rehabilitation through engaging gameplay.

The system was developed using Unity 2020.1.17 game engine (https://unity.com/), selected for its flexibility, performance and cross-platform support. Three-dimensional modelling, texturing and animation was done with

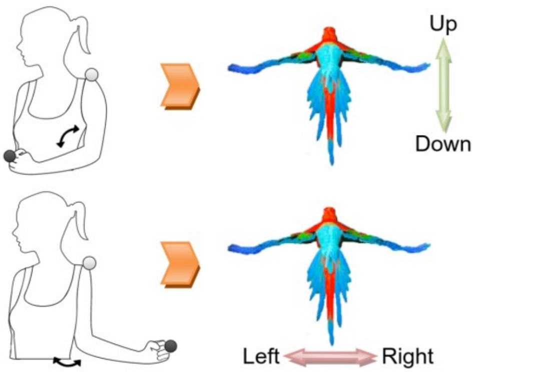


Fig. 2. Illustration of the character control mechanism in the exergame. Up and down movements are controlled by elbow flexion and extension, while left and right movements are controlled by shoulder internal and external rotation.

Fig. 3. Exergame scenario.

Blender 2.79 (https://www.blender.org/), a free and widely adopted tool. The 2D vectors were created using the software GIMP – GNU Image Manipulation Program, version 2.10.36 (https://www.gimp.org/), a free alternative to commercial image editors. The OpenCV-Unity integration was done using the OpenCV plus Unity 1.7.1, a free plugin [35]. OpenCV – Open Computer Vision Library 4.6 (https://opencv.org/) was used for camera calibration, image segmentation, and 3D reconstruction. It was chosen for its documentation and good performance.

The application was developed and tested on a laptop equipped with an Intel Core i7 processor, 32 GB of RAM, and an NVIDIA GTX 1660 Ti GPU, ensuring real-time execution at approximately 30 frames per second.

The system was designed for autonomous use by patients, with optional therapist support. A training video guides users through setup, calibration, and gameplay. Initial calibration includes camera distortion correction using checkerboard images and color calibration to define sphere color, diameter, and forearm length. This enables real-time 3D elbow angle estimation via image segmentation and geometric reconstruction.

The exergame is set in a beach environment, where elbow flexion controls vertical bird movement and shoulder rotation controls horizontal movement (Fig. 2). The character control was designed to encourage repeated elbow flexion and extension, which are fundamental movements in upper limb rehabilitation. Additionally, the game incorporates cognitively engaging tasks to stimulate attention, planning, and reaction time, all within a motivating and playful context. The objective is to collect virtual fruits, with performance metrics displayed in real time (Fig. 3).

At the end of the session, a motor coordination test requires tracing an on-screen pattern with the sphere. After completion, the system generates a report and sends it via email. All data is stored in the cloud, enabling remote monitoring and progress tracking by therapists.

B. Elbow Angle Estimation Pipeline

1) Camera Calibration: To mitigate lens distortion, intrinsic and extrinsic parameters of the camera were obtained using OpenCV's implementation of the Brown–Conrady distortion model [25]. A checkerboard pattern with 10 × 7 squares (23.3 mm) was printed and captured from 50 different angles. The model corrects both radial and tangential distortions:

- Radial distortion:

$$x_{corr} = x\left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right)$$
$$y_{corr} = y\left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right)$$

- Tangential distortion:

$$x_{corr} = x + \left[2p_1 xy + p_2\left(r^2 + 2x^2\right)\right]$$
$$y_{corr} = y + \left[p_1\left(r^2 + 2y^2\right) + 2p_2 xy\right]$$

where $r^2 = x^2 + y^2$ and k_1, k_2, k_3, p_1, p_2 are the distortion coefficients.

2) Marker Segmentation: The system uses two colored spheres (4 cm diameter) attached to the arm, one on the shoulder and the other on hand. Frames are preprocessed with blurring, morphological filtering (erosion and dilation), and HSV-based color thresholding. Contours are identified and filtered based on circularity and size constraints. The centroid of each contour is extracted and stored as the 2D marker position in pixel space.

3) 3D Reconstruction via Perspective-n-Point: Given the camera's intrinsic matrix and the known radius and position of the spheres in real-world coordinates, their 3D positions (X, Y, Z) were retrieved using the Perspective-n-Point algorithm [36].

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathrm{K}\,[\mathrm{R}|\,\mathrm{t}]\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

where K is the camera matrix, and R, t are rotation and translation vectors.

4) Elbow Angle Estimation: Using 3D positions of the shoulder and the hand spheres $\vec{S}$ and $\vec{H}$, the Euclidean distance $d = \left\|\vec{S} - \vec{H}\right\|$ was computed. Assuming known distances from elbow to each sphere (d_1 and d_2)—measured and provided during the configuration stage, and computing the third side d_3 (distance from shoulder to hand), the elbow angle θ was

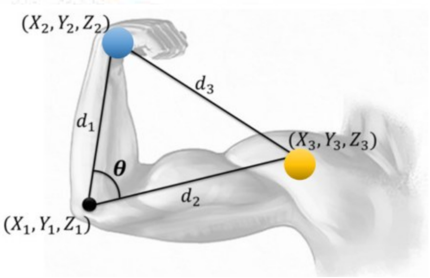


Fig. 4. The elbow angle (θ) is computed at the vertex formed by the three-dimensional points corresponding to the shoulder (X_1, Y_1, Z_1), elbow (X_2, Y_2, Z_2) and hand (X_3, Y_3, Z_3), using distances d_1, d_2 and d_3 as triangle sides.

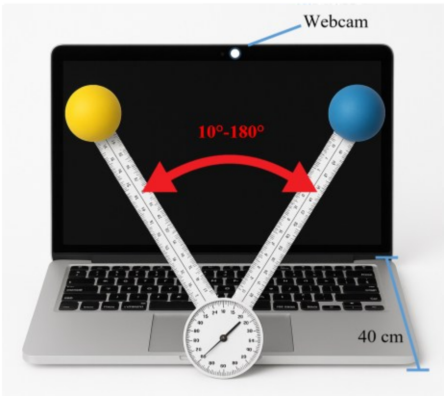


Fig. 5. Representation of the goniometer with spheres attached for the angle value test.

derived from the Law of Cosines:

$$\theta = \cos^{-1}\left(\frac{d_1^2 + d_2^2 - d_3^2}{2d_1d_2}\right)$$

This method allows accurate angle estimation in 3D space under free arm movement (Fig. 4).

C. Evaluation Protocols

1) Accuracy of Joint Angle Estimation: To assess the accuracy of the elbow joint angles estimated by the system, we conducted a controlled validation procedure using a manual goniometer as the reference instrument. This device is widely considered the clinical gold standard for joint angle measurement. As illustrated in Fig. 5, two 4 cm-diameter polystyrene spheres of different colors were affixed to the distal ends of the goniometer arms to emulate anatomical landmarks of a human upper limb.

The goniometer was placed on a vertical plane, aligned parallel to the webcam of a standard laptop positioned 40 cm away. The webcam (HD 1280 × 720 resolution) had been previously calibrated with 50 reference images to correct for lens distortion and improve 3D triangulation fidelity. During the test, the goniometer arms were manually adjusted at 10° increments, spanning a range from 10° to 180°, and held static at each target angle while 200 samples were acquired. This value was empirically chosen as the highest round number that allowed real-time acquisition.

The estimated angles were computed in real time, and the distribution of values at each reference position was recorded. For each angle, the arithmetic mean of the 200 estimated values was used for comparison against the goniometer reference. The angular estimation error E_θ was calculated as the absolute difference between the estimated and reference values, as defined by:

$$E_\theta = \left|\theta_{goniometer} - \theta_{system}\right|,$$

where E_θ is expressed in degrees. All recordings were conducted in a controlled indoor environment with stable artificial lighting to minimize external interferences.

2) Camera Calibration Robustness Analysis: To investigate the impact of the number of calibration images on system accuracy, seven calibration sets were assembled, containing 46, 25, 10, 5, 3, 2, and 1 image(s), respectively. These values were selected to simulate varying levels of calibration quality. Each set was used independently to perform the camera calibration described in Section II.

Following each calibration procedure, the goniometer was positioned at a fixed angle of 60°, simulating a static joint configuration. The same test setup described in the previous test was used, maintaining the camera at a distance of 40 cm from the goniometer, with the colored spheres approximately centered in the image frame. For each calibration setting, 200 angle estimations were captured under identical environmental conditions.

The standard deviation of the estimated angles for each calibration set was computed as a measure of stability and repeatability. In parallel, the reprojection error reported by OpenCV was recorded, providing an indirect metric of calibration accuracy. Both metrics were plotted and analyzed to determine the minimum number of images required to ensure reliable system performance without unnecessary computational overhead.

3) Evaluation of Spatial Parameters on Measurement Accuracy: To investigate how variations in marker size and camera distance influence angular estimation accuracy, a systematic test was performed using five different diameters of colored polystyrene spheres (3.0 cm, 3.5 cm, 4.0 cm, 5.0 cm, and 6.0 cm). For each configuration, a pair of spheres of the same diameter was affixed to the ends of a manual goniometer, which was maintained at a constant angle of 60° throughout the experiment.

The system was evaluated at four distances between the camera and the goniometer along the z-axis: 40 cm, 80 cm, 120 cm, and 160 cm. For each combination of sphere diameter and camera distance, 100 angular estimations were acquired under fixed lighting conditions. The webcam and calibration setup were the same as described previously.

For each configuration, the mean estimated angle and its 95% confidence interval were calculated. To evaluate accuracy, the absolute angular error was computed sample by sample as the absolute difference between each estimated angle and the reference value of 60°. The mean absolute angular error for each condition was then calculated across the 100 samples. A one-way ANOVA was performed separately for each distance to assess whether differences in mean angular error across the five sphere diameters were statistically significant. The results were analyzed to identify which combinations yielded the lowest angular errors, providing insight into the optimal spatial conditions for system deployment in real-world rehabilitation scenarios.

4) Angular Error Mapping Across Image Coordinates: This test investigated the impact of the spatial location of the tracked markers (spheres) within the image frame on the accuracy of elbow angle estimation. Even after camera calibration, peripheral image areas tend to exhibit residual optical distortions, which may introduce systematic errors in measurements.

The setup consisted of a goniometer fixed at 60°, with two colored polystyrene spheres (4 cm diameter) affixed to its arms. A webcam (1280 × 720 px resolution), previously calibrated with 50 images, was placed 40 cm in front of the device.

To isolate the effect of horizontal and vertical displacements, the test was split into two components:

X-axis variation: the spheres were translated laterally across the width of the image, from left to right, while maintaining a constant Y-position. The horizontal coordinate of the rightmost sphere was recorded, and the corresponding error was noted as E_x. 250 samples were recorded.

Y-axis variation: similarly, the spheres were moved vertically from top to bottom, keeping the X-position fixed, with vertical coordinate y recorded and respective error E_y. 250 samples were recorded.

The absolute angular error for each position was directly recorded by the system and used to analyze the influence of spatial location. The absolute angular error was defined as: $E_\theta = \left|\theta_{system} - 60°\right|$.

Each set of measurements was modeled using a second-degree polynomial regression

$$E_x(x) = a_1x^2 + a_2x + a_3$$
$$E_y(y) = b_1y^2 + b_2y + b_3$$

where $E_x(x)$ and $E_y(y)$ represent the angular error as a function of the horizontal and vertical coordinates, respectively.

5) Degradation of Angular Estimation Accuracy Under Dynamic Conditions: To investigate how the movement speed of the spheres affects the angular accuracy, a test was performed with the goniometer fixed at an angle of 60°, using two spheres of 4 cm in diameter, one on each arm. The goniometer was manually moved up and down along the vertical (Y) axis, introducing a range of movement speeds, until 250 values were captured.

The Y-velocity of the spheres was calculated in cm/s using frame-by-frame position differences captured by the camera. The corresponding angular error was calculated for each frame as the absolute difference from the target angle. The Pearson correlation coefficient r between velocity v_i and angular error e_i was computed using:

$$r = \frac{\sum (v_i - \overline{v})(e_i - \overline{e})}{\sqrt{\sum (v_i - \overline{v})^2 \sum (e_i - \overline{e})^2}}$$

where $\overline{v}$ and $\overline{e}$ represent the means of velocity and error, respectively.

A linear regression model was also fitted to predict the angular error based on sphere velocity. An ANOVA test was conducted across the velocity ranges to assess whether differences in angular error were statistically significant.

6) Validation Against a Deep Learning-Based Pose Estimation Framework: To evaluate the accuracy of the proposed system under conditions that challenge AI-based pose estimation, we compared it with a state-of-the-art deep learning method, MediaPipe Pose Landmarker (Full Body) [31]. A subject was instructed to maintain the right elbow flexed at a fixed angle of 90°, while the shoulder joint was slowly rotated from left to right, sweeping the arm across the camera's field of view. This motion maintained the elbow at a constant angle, but progressively altered the spatial positioning and silhouette visibility.

To track joint positions, a 4 cm blue sphere was affixed to the right shoulder and a 4 cm yellow sphere to the right hand. The marker trajectory was recorded using a 720p webcam positioned in front of the subject. The test took 7 seconds at 30 frames per second. The elbow angle was estimated throughout the trajectory using both (1) the proposed marker-based method, which applies geometric projection and perspective correction based on the tracked spheres, and (2) MediaPipe, which estimates elbow angles from 3D keypoints inferred through skeletal pose detection.

Both methods were applied to the exact same video footage. A Python script was used to compute the elbow angle from the MediaPipe landmarks based on the angle formed by the shoulder–elbow–hand keypoints. Prior validation ensured that the presence of the spheres did not interfere with MediaPipe detection.

7) Usability Evaluation With Stroke Patients: To assess the perceived usability of the proposed system in a clinical setting, a usability study was conducted involving eight post-stroke patients. The System Usability Scale (SUS) [37] was employed as a standardized instrument to quantify user experience and interface satisfaction. Each patient answered 10 SUS items on a 5-point Likert scale, ranging from "strongly disagree" to "strongly agree." These items evaluate key usability factors such as learnability, efficiency, confidence, and system complexity.

Participants were selected based on the following inclusion criteria: chronic stroke ($\geq$ 6 months post-onset), mild to moderate motor or cognitive impairment, and adequate visual and auditory function to interact with the system. Patients with severe cognitive deficits, communication barriers, or uncontrolled comorbidities were excluded. Each participant underwent supervised testing under the guidance of a licensed physical therapist, and informed consent was obtained prior to the sessions. Ethics approval was granted by the Federal University of Uberlandia (protocol no. 39232820.2.0000.5152).

The individual SUS item scores were plotted using a radar chart to visualize user-specific perceptions across usability dimensions. The final SUS score was computed as the mean of individual scores converted to a 0–100 scale using the standard SUS algorithm.

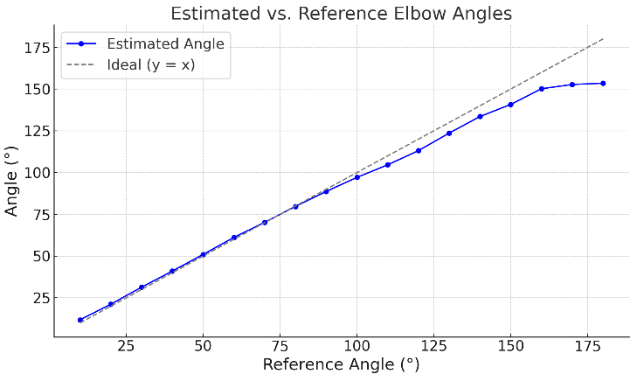


Fig. 6. Comparison between reference angles obtained using a manual goniometer and the mean estimated angles produced by the system across the tested range (10°–180°). The dashed line represents the ideal estimation (y = x). The solid line shows the mean values obtained by the system. A deviation from the ideal trend is observed for angles above 110°, indicating reduced accuracy in estimating wide joint openings.

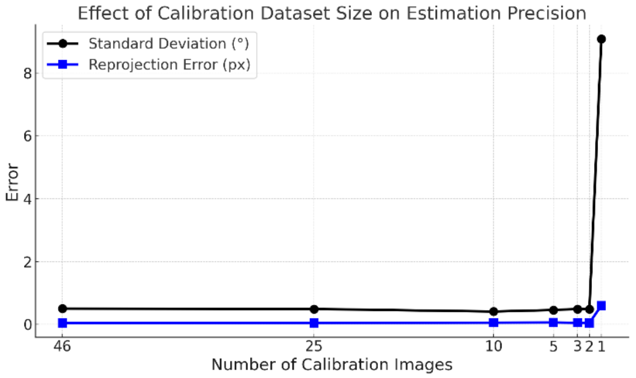


Fig. 7. Effect of calibration dataset size on system precision. The graph presents the standard deviation of estimated angles and the reprojection error as a function of the number of images used during the calibration process. A sharp degradation is observed with only one image.

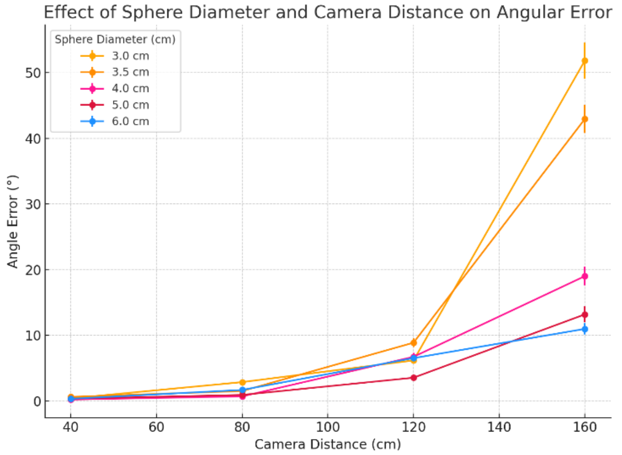


Fig. 8. Effect of sphere diameter and camera distance on angular estimation error. Each line represents a different sphere diameter, and error bars indicate 95% confidence intervals for the mean absolute error. A consistent increase in angular error is observed with greater camera distance. The lowest errors were observed for distances of 40–80 cm.

III. RESULTS

This section presents the experimental results obtained to evaluate the performance of the proposed vision-based system for elbow angle estimation during rehabilitation tasks.

A. Accuracy of Joint Angle Estimation

To evaluate the reliability of the proposed system in estimating elbow joint angles, a controlled benchmark test was conducted using a manual goniometer as reference. The test setup and procedure were described in Section II. The results, based on 200 samples for each reference angle from 10° to 180°, are presented in Fig. 6 and Table II.

Fig. 6 illustrates the relationship between the system's estimated angles and the true reference angles. The ideal scenario (where estimated = reference) is represented by a dashed line. The system performs with high accuracy for angles up to approximately 110°, closely following the ideal curve. However, for obtuse angles, particularly beyond 120°, the estimates increasingly deviate from the reference, revealing the method's limitations when the inter-sphere distance varies slowly with respect to angle.

Table II summarizes the estimation performance in detail, for each angle. To aid interpretation, the "Absolute Error (°)" column is color-coded: green for low error (≤5°), yellow for moderate error (5°–10°), and red for high error (>10°). These results suggest that the system is best suited for applications involving flexion ranges up to 110°, where angular precision remains within clinically acceptable limits.

B. Camera Calibration Robustness Analysis

To evaluate the effect of calibration dataset size on angular estimation precision, the system was tested under identical conditions using calibration sets with varying numbers of images. As shown in Fig. 7, both the reprojection error (in pixels) and the standard deviation of angle estimations (in degrees) remained low and stable when using two or more calibration images. However, a substantial increase in angular variability was observed when calibration was performed with only a single image, indicating poor geometric correction.

This result highlights that one image is insufficient to capture the necessary distortion parameters for reliable camera calibration. In contrast, the use of two calibration images was found to be sufficient under the test conditions, yielding comparable precision to larger datasets. No significant improvements were observed when increasing the dataset beyond this threshold, suggesting that the system is robust to calibration size within this range.

C. Evaluation of Spatial Parameters on Measurement Accuracy

This experiment investigated the influence of spherical marker size and camera distance on the accuracy of elbow angle estimation. The objective was to determine the optimal combination of parameters for minimizing angular error. Five different marker diameters (3.0 cm, 3.5 cm, 4.0 cm, 5.0 cm, and 6.0 cm) were tested at four distances from the camera (40 cm, 80 cm, 120 cm, and 160 cm), while maintaining a fixed goniometer angle of 60°.

The results are illustrated in Fig. 8. Angular error increased consistently with greater camera distance across all sphere sizes. The ten configurations with the lowest errors occurred at the shortest distances tested (40 and 80 cm), with each sphere size represented at least twice among these optimal results. Notably, larger markers ($\geq$ 4.0 cm) demonstrated more stable performance over increasing distances, while smaller spheres showed a sharp rise in error at long ranges.

A one-way ANOVA conducted for each distance confirmed significant differences in angular error between marker diameters ($p < 0.001$ at all distances).

These findings suggest that optimal system performance is achieved when the camera is placed within 80 cm of the markers, and that spheres ranging from 3.5 cm to 6.0 cm are suitable for maintaining acceptable accuracy in clinical applications.

D. Angular Error Mapping Across Image Coordinates

The regression analysis revealed that both horizontal and vertical positions influence the angular estimation accuracy. Specifically, the quadratic fit showed that the error is minimized near the center of the image and increases towards the borders, especially at the upper-left and lower-right corners. This behavior is consistent with the expected distortion profile from common consumer-grade lenses. The resulting equations were:

- Horizontal (X axis):

$$E_x(x) = 2.52 \times 10^{-5} x^2 - 0.03x + 12.02$$

- Vertical (Y axis)

$$E_y(y) = 4.16 \times 10^{-5} y^2 - 0.03y + 5.22$$

These models were derived using least squares fitting and show a clear parabolic trend for both axes.

The heatmaps (Fig. 9) illustrate the spatial variability of angular estimation accuracy across the image frame:

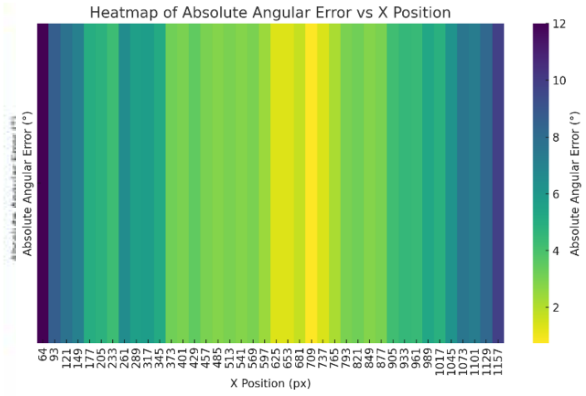
- Fig. 9a shows that the angular error along the X-axis is minimal near the image center ($x \approx 640$ px) and increases as the spheres approach the left or right edges.
- Fig. 9b reveals a similar pattern for the Y-axis, with lower errors in the mid-height of the frame and higher errors near the top and bottom boundaries.

These results align with the parabolic regression trends and indicate that lens distortion and non-uniform calibration affect peripheral areas more significantly. Therefore, patients should ideally perform movements near the center of the image frame for improved tracking accuracy.

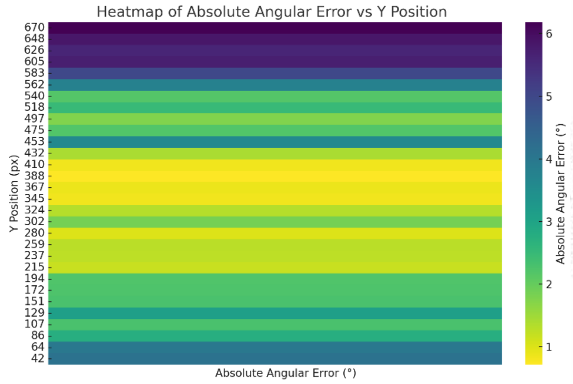
E. Degradation of Angular Estimation Accuracy Under Dynamic Conditions

The impact of sphere velocity on angular measurement accuracy is illustrated in Fig. 10. Angular error remained minimal below 10 cm/s and increased progressively at higher speeds.

A Pearson correlation coefficient of $r = 0.70$ was obtained, indicating a strong positive correlation between speed and error. A linear regression model further demonstrated this relationship, with the equation: $E_\theta = 0.52v + 1.86$.



(a)



(b)

Fig. 9. a. Heatmap of absolute angular error as a function of horizontal (X) position in the image. Lighter regions represent lower error. b. Heatmap of absolute angular error as a function of vertical (Y) position in the image. Lighter regions represent lower error.

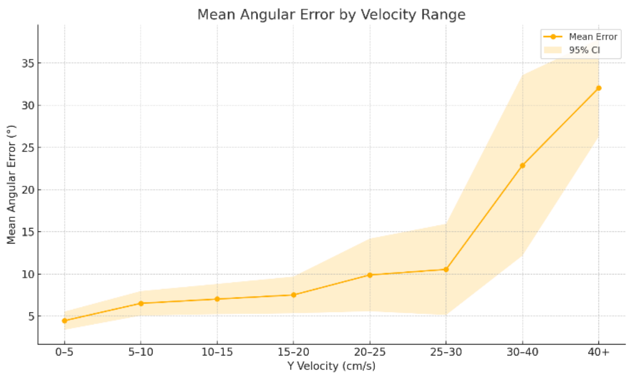


Fig. 10. Mean angular error by velocity range, with 95% confidence intervals (yellow area). A sharp increase in error confirms a statistically significant degradation in tracking accuracy at higher velocities (ANOVA, $p < 0.001$).

An ANOVA test performed across velocity groups confirmed that the differences in angular error between speed ranges were statistically significant ($p < 0.001$), reinforcing the importance of movement speed in tracking precision. These findings indicate that, to ensure accurate measurements, users should perform movements at moderate and controlled speeds.

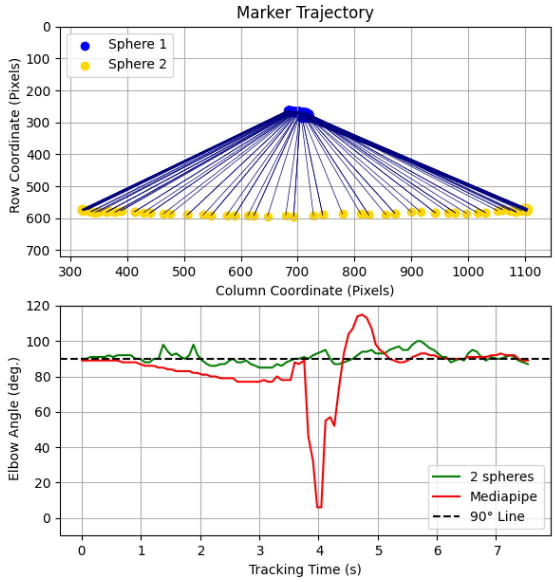


Fig. 11. Comparison of elbow angle estimation between the proposed marker-based system and MediaPipe. The top panel illustrates the 2D pixel trajectories of the two tracked spheres. The bottom panel shows angle estimates over time. While the proposed method remains stable around 90°, MediaPipe presents significant errors when the arm moves close to the torso.

F. Validation Against a Deep Learning-Based Pose Estimation Framework

The comparison between the proposed marker-based solution and MediaPipe is illustrated in Fig. 11. The upper subplot shows the pixel-space trajectories of both spheres as the arm moved from left to right. The lower subplot presents the estimated elbow angles across time.

The proposed method (green curve) maintained stable estimations near the expected 90° line (dashed black), even when the arm approached the torso, demonstrating resilience to occlusion and silhouette blending. In contrast, MediaPipe (red curve) exhibited substantial error at the center of the trajectory, particularly between 3 and 5 seconds, where angle estimates dropped below 50°. This discrepancy is likely caused by MediaPipe's reduced ability to distinguish arm landmarks when the arm is adjacent to the body, leading to landmark misplacement and large angular errors.

G. Usability With Patients

Fig. 12 shows the System Usability Scale (SUS) [37] responses for each of the eight participants. The radar chart illustrates individual ratings across 10 core usability items. Most patients reported high agreement with statements indicating ease of use, confidence, and system integration. Conversely, statements such as "need for technical support" and "need to learn a lot to use" received low agreement, suggesting that some participants may need support to use the system.

The average SUS score across all patients was 92.5, which classifies the system in the "Excellent" usability range. These results suggest that the system is well-accepted by users with chronic post-stroke conditions and may serve as a feasible tool for telerehabilitation. Detailed patient characteristics are presented in Table I. A real gameplay session with a post-stroke volunteer is shown in Fig. 13.

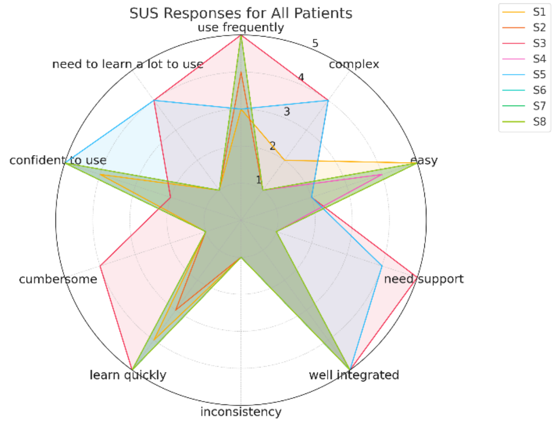


Fig. 12. Radar chart displaying the SUS responses from eight stroke patients across the 10 standard usability dimensions.

TABLE I
DETAILS OF THE STROKE SUBJECTS

Subject	Age	Years since onset	Gender	FMA[1]
S1	33	5	F	62
S2	46	8	M	43
S3	62	32	F	21
S4	52	8	M	12
S5	49	5	F	26
S6	36	2	M	15
S7	45	13	M	26
S8	37	1	M	48

[1] FMA refers to the Fugl-Meyer Assessment Upper Extremity, a score between 0 (no function) and 66 (intact).

TABLE II
ERROR OF THE ESTIMATED ANGLES

Reference Angle (°)	Mean Estimated Angle (°)	95% Confidence Interval (°)	Absolute Error (°)[1]	Relative Error (%)
10	11.85	±0.06	1.85	18.46
20	21.17	±0.05	1.17	5.87
30	31.38	±0.07	1.38	4.59
40	41.01	±0.06	1.01	2.52
50	50.99	±0.06	0.99	1.98
60	61.16	±0.08	1.16	1.93
70	70.19	±0.08	0.19	0.27
80	79.74	±0.09	0.26	0.32
90	88.68	±0.10	1.32	1.47
100	97.11	±0.13	2.89	2.89
110	104.65	±0.16	5.35	4.86
120	113.12	±0.17	6.88	5.73
130	123.65	±0.25	6.35	4.89
140	133.56	±0.23	6.44	4.59
150	140.80	±0.37	9.20	6.13
160	150.19	±0.51	9.81	6.13
170	152.75	±0.46	17.25	10.15
180	153.54	±0.63	26.46	14.70

[1]Color coding in the "Absolute Error (°)" column indicates the magnitude of error for each reference angle: green (≤5°) for high accuracy, yellow (5°–10°) for moderate deviation, and red (>10°) for significant error. This visual indicator highlights the system's performance degradation for large, obtuse angles.

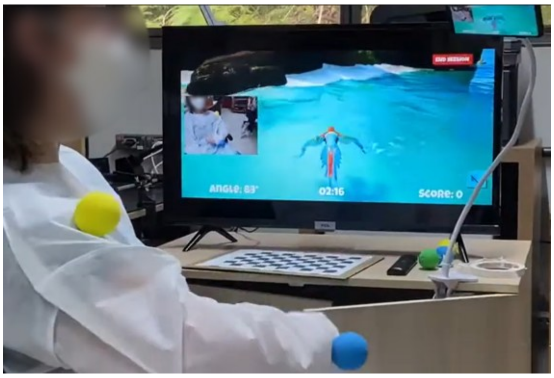


Fig. 13. Pilot test of the proposed exergame system with a post-stroke patient. The game runs on a smartphone and is mirrored on a television screen. Colored markers on the arm enable real-time 3D tracking using the smartphone's webcam.

IV. Discussion

The proposed marker-based elbow tracking system demonstrates a promising solution for accessible post-stroke rehabilitation, leveraging low-cost vision-based techniques without the need for wearable electronics or depth sensors. A series of technical experiments provided a comprehensive evaluation of its performance under various operational conditions. The results offer insights into its applicability in real-world clinical scenarios.

The system achieved high accuracy in joint angle estimation for flexion ranges below 110°, with absolute errors under 5°, which is within clinically acceptable thresholds [38]. Although accuracy declined for angles above 110°, this range exceeds most movements commonly targeted in stroke rehabilitation. Thus, the limitation is unlikely to affect core clinical applications.

Environmental and operational factors were also explored. Camera calibration was found to be robust even with as few as two calibration images, while larger datasets did not yield significant gains. Additionally, spatial configuration analyses showed that camera proximity ($\leq$80 cm) and marker diameters $\geq$3.5 cm optimize measurement reliability, aligning with principles of image resolution and object segmentation.

Another critical finding involves the influence of spatial position within the image frame. Even after lens distortion correction, regression models revealed a parabolic distribution of error along both the horizontal and vertical axes, with the lowest errors occurring at the image center. This suggests that patients should ideally operate within central camera regions to minimize systematic distortions.

Movement velocity emerged as a significant variable. A Pearson correlation coefficient of $r = 0.70$ was observed between marker velocity and angular error, indicating a strong positive relationship. Linear regression analysis confirmed this trend, with angular errors remaining below 5° for velocities under 10 cm/s, increasing to approximately 10° in the 20–30 cm/s range, and exceeding 30° at speeds above 40 cm/s. These results demonstrate that faster motion leads to increased error, primarily due to motion blur inherent to consumer-grade webcams. This reinforces the recommendation for slow, controlled movements during gameplay to ensure reliability.

The system's performance was also benchmarked against MediaPipe, a state-of-the-art deep learning framework for 3D pose estimation. While MediaPipe excelled under clear limb configurations, it exhibited substantial errors under occlusion, particularly when the arm was close to the torso. The proposed method, by contrast, maintained stable accuracy across a broader range of spatial poses, a key advantage in rehabilitation scenarios where limb occlusion is common.

Finally, usability testing with eight post-stroke patients yielded a System Usability Scale (SUS) score of 92.5, which is considered "excellent." Patients generally rated the system as easy to use, well-integrated, and confidence-inspiring. However, some required therapist assistance, indicating that while the system is accessible, further refinement is needed to enhance autonomy for all users.

Finally, while the system is designed to run on a variety of devices, optimal performance is achieved with a device featuring a dedicated GPU, at least 8 GB of RAM, and an HD (720p or higher) webcam. Improved accuracy may be observed with external webcams offering better optics, but such upgrades are optional and not required for the system to function reliably.

The anonymized dataset generated and analysed during the current study is available from the corresponding author upon reasonable request. The source code is not publicly available.

V. Conclusion

This study presented the development and comprehensive evaluation of a low-cost, vision-based 3D elbow tracking system designed for post-stroke rehabilitation. The system's unique marker-based tracking method, combined with geometric angle estimation and real-time exergame interaction, enables effective and accessible rehabilitation without the need for specialized hardware.

The proposed solution demonstrated high accuracy within clinically relevant flexion ranges and robust performance under various calibration, spatial, and dynamic conditions. Its performance surpassed that of AI-based skeletal pose estimation (MediaPipe) under occluded limb scenarios, highlighting its suitability for rehabilitation tasks where precision is essential.

The system was well-received by post-stroke patients, achieving a SUS score indicative of high usability. Although minor limitations were observed, such as increased error at high velocities and peripheral image locations, the results affirm the system's feasibility for home-based or supervised clinical use.

Future work will involve comparative validation against gold-standard systems like Vicon, focusing on metrics such as dynamic tracking accuracy, test-retest reliability and angular resolution across movement ranges. In addition, telehealth-oriented features will be explored. Longitudinal clinical trials are also planned to evaluate long-term rehabilitation outcomes and assess patient engagement over extended use. Overall, the system represents a scalable, cost-effective alternative to sensor-based rehabilitation, promoting broader access to therapy and enabling quantitative tracking of motor recovery.

ACKNOWLEDGMENT

The Article Processing Charge (APC) for the publication of this research was funded by the Coordination for the Improvement of Higher Education Personnel–CAPES (ROR ID: 00x0ma614). For the purpose of open access, the authors have assigned a Creative Commons CC BY license to any accepted author manuscript version arising from this submission.

REFERENCES

[1] C. W. Tsao et al., "Heart disease and stroke Statistics–2022 update: A report from the American heart association," *Circulation*, vol. 145, no. 8, pp. e153–e639, Feb. 2022, doi: 10.1161/cir.0000000000001052.

[2] V. Feigin et al., "Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016," *New England J. Med.*, vol. 379, pp. 2429–2437, Dec. 2018, doi: 10.1056/NEJMoa1804492.

[3] A. Heller, D. T. Wade, V. A. Wood, A. Sunderland, R. L. Hewer, and E. Ward, "Arm function after stroke: Measurement and recovery over the first three months," *J. Neurol., Neurosurgery Psychiatry*, vol. 50, no. 6, pp. 714–719, Jun. 1987, doi: 10.1136/jnnp.50.6.714.

[4] A. Sunderland, D. Tinson, L. Bradley, and R. L. Hewer, "Arm function after stroke. An evaluation of grip strength as a measure of recovery and a prognostic indicator," *J. Neurol., Neurosurgery Psychiatry*, vol. 52, no. 11, pp. 1267–1272, Nov. 1989, doi: 10.1136/jnnp.52.11.1267.

[5] H. Nakayama, H. S. Jørgensen, H. Otto Raaschou, and T. Skyhøj Olsen, "Recovery of upper extremity function in stroke patients: The Copenhagen stroke study," *Arch. Phys. Med. Rehabil.*, vol. 75, no. 4, pp. 394–398, Apr. 1994, doi: 10.1016/0003-9993(94)90161-9.

[6] G. Saposnik and M. Levin, "Virtual reality in stroke rehabilitation: A meta-analysis and implications for clinicians," *Stroke*, vol. 42, no. 5, pp. 1380–1386, May 2011, doi: 10.1161/strokeaha.110.605451.

[7] D. T. Wade, R. Langton-Hewer, V. A. Wood, C. E. Skilbeck, and H. M. Ismail, "The hemiplegic arm after stroke: Measurement and recovery," *J. Neurol., Neurosurgery Psychiatry*, vol. 46, no. 6, pp. 521–524, Jun. 1983, doi: 10.1136/jnnp.46.6.521.

[8] A. Pyae, M. Luimula, and J. Smed, "Rehabilitative games for stroke patients," *EAI Endorsed Trans. Serious Games*, vol. 1, no. 4, p. e2, 2015.

[9] K. Thomson, A. Pollock, C. Bugge, and M. C. Brady, "Commercial gaming devices for stroke upper limb rehabilitation: A survey of current practice," *Disab. Rehabil., Assistive Technol.*, pp. 1–8, Jan. 2015, doi: 10.3109/17483107.2015.1005031.

[10] D. Jack et al., "Virtual reality-enhanced stroke rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 9, no. 3, pp. 308–318, Mar. 2001.

[11] WIi. *Wii | Nintendo*. Accessed: Apr. 23, 2024. [Online]. Available: https://www.nintendo.pt/Wii/Wii-94559.html

[12] Microsoft. *Azure Kinect DK*. Accessed: Apr. 23, 2024. [Online]. Available: https://azure.microsoft.com/pt-br/products/Kinect-dk

[13] B.-M. Gutiérrez-Pérez, A.-V. Martín-García, A. Murciano-Hueso, and A.-P. de Oliveira Cardoso, "Use of serious games with older adults: Systematic literature review," *Humanities Social Sci. Commun.*, vol. 10, no. 1, p. 939, Dec. 2023, doi: 10.1057/s41599-023-02432-0.

[14] J. Z. Nie, J. W. Nie, N.-T. Hung, R. J. Cotton, and M. W. Slutzky, "Portable, open-source solutions for estimating wrist position during reaching in people with stroke," *Sci. Rep.*, vol. 11, no. 1, p. 22491, Nov. 2021, doi: 10.1038/s41598-021-01805-2.

[15] D. Quintana, A. Rodríguez, and I. Boada, "Limitations and solutions of low cost virtual reality mirror therapy for post-stroke patients," *Sci. Rep.*, vol. 13, no. 1, p. 14780, Sep. 2023, doi: 10.1038/s41598-023-40546-2.

[16] S. M. Parker, B. Ricks, J. Zuniga, and B. A. Knarr, "Comparison of virtual reality to physical box and blocks on cortical an neuromuscualar activations in young adults," *Sci. Rep.*, vol. 13, no. 1, p. 16567, Oct. 2023, doi: 10.1038/s41598-023-43073-2.

[17] G. Abbate, A. Giusti, L. Randazzo, and A. Paolillo, "A mirror therapy system using virtual reality and an actuated exoskeleton for the recovery of hand motor impairments: A study of acceptability, usability, and embodiment," *Sci. Rep.*, vol. 13, no. 1, p. 22881, Dec. 2023, doi: 10.1038/s41598-023-49571-7.

[18] J. S. Lora-Millan, F. J. Sanchez-Cuesta, J. P. Romero, J. C. Moreno, and E. Rocon, "Robotic exoskeleton embodiment in post-stroke hemiparetic patients: An experimental study about the integration of the assistance provided by the REFLEX knee exoskeleton," *Sci. Rep.*, vol. 13, no. 1, p. 22908, Dec. 2023, doi: 10.1038/s41598-023-50387-8.

[19] F. Noveletto et al., "Biomedical serious game system for balance rehabilitation of hemiparetic stroke patients," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2179–2188, Nov. 2018, doi: 10.1109/TNSRE.2018.2876670.

[20] M. Ghassemi et al., "Development of an EMG-controlled serious game for rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 2, pp. 283–292, Feb. 2019, doi: 10.1109/TNSRE.2019.2894102.

[21] Y. Jiang, Z. Liu, T. Liu, M. Ma, M. Tang, and Y. Chai, "A serious game system for upper limb motor function assessment of hemiparetic stroke patients," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 2640–2653, 2023, doi: 10.1109/TNSRE.2023.3281408.

[22] J. Lee, J. E. Park, B. H. Kang, and S. N. Yang, "Efficiency of botulinum toxin injection into the arm on postural balance and gait after stroke," *Sci. Rep.*, vol. 13, no. 1, p. 8426, May 2023, doi: 10.1038/s41598-023-35562-1.

[23] R. de-la-Torre, E. D. Oña, J. G. Victores, and A. Jardón, "SpasticSim: A synthetic data generation method for upper limb spasticity modelling in neurorehabilitation," *Sci. Rep.*, vol. 14, no. 1, p. 1646, Jan. 2024, doi: 10.1038/s41598-024-51993-w.

[24] F. Al Farid et al., "A structured and methodological review on vision-based hand gesture recognition system," *J. Imag.*, vol. 8, no. 6, p. 153, May 2022, doi: 10.3390/jimaging8060153.

[25] T. Takebayashi, Y. Uchiyama, Y. Okita, and K. Domen, "Development of a program to determine optimal settings for robot-assisted rehabilitation of the post-stroke paretic upper extremity: A simulation study," *Sci. Rep.*, vol. 13, no. 1, p. 9217, Jun. 2023, doi: 10.1038/s41598-023-34556-3.

[26] L. Bai, M. G. Pepper, Y. Yan, S. K. Spurgeon, M. Sakel, and M. Phillips, "Quantitative assessment of upper limb motion in neurorehabilitation utilizing inertial sensors," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 2, pp. 232–243, Mar. 2015, doi: 10.1109/TNSRE.2014.2369740.

[27] L. Ding et al., "Camera-based mirror visual feedback: Potential to improve motor preparation in stroke patients," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 9, pp. 1897–1905, Sep. 2018, doi: 10.1109/TNSRE.2018.2864990.

[28] Y. Gu et al., "A review of hand function rehabilitation systems based on hand motion recognition devices and artificial intelligence," *Brain Sci.*, vol. 12, no. 8, p. 1079, 2022, doi: 10.3390/brainsci12081079.

[29] A. Abedi, T. J. F. Colella, M. Pakosh, and S. S. Khan, "Artificial intelligence-driven virtual rehabilitation for people living in the community: A scoping review," *npj Digit. Med.*, vol. 7, no. 1, p. 25, Feb. 2024, doi: 10.1038/s41746-024-00998-w.

[30] S. Rahman, S. Sarker, A. K. M. N. Haque, M. M. Uttsha, M. F. Islam, and S. Deb, "AI-driven stroke rehabilitation systems and assessment: A systematic review," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 192–207, 2023, doi: 10.1109/TNSRE.2022.3219085.

[31] Google. *MediaPipe | Google for Developers*. Accessed: Jan. 21, 2024. [Online]. Available: https://developers.google.com/mediapipe

[32] L. C. Klein et al., "Assessing the reliability of AI-based angle detection for shoulder and elbow rehabilitation," in *Proc. 3rd Int. Conf. Optim.*, 2024, pp. 3–18, doi: 10.1007/978-3-031-53036-4_1.

[33] P. Palani, S. Panigrahi, S. A. Jammi, and A. Thondiyath, "Real-time joint angle estimation using mediapipe framework and inertial sensors," in *Proc. IEEE 22nd Int. Conf. Bioinf. Bioengineering (BIBE)*, Nov. 2022, pp. 128–133, doi: 10.1109/BIBE55377.2022.00035.

[34] J.-Y. Choi, E. Ha, M. Son, J.-H. Jeon, and J.-W. Kim, "Human joint angle estimation using deep learning-based three-dimensional human pose estimation for application in a real environment," *Sensors*, vol. 24, no. 12, p. 3823, Jun. 2024, doi: 10.3390/s24123823.

[35] Unity. *OpenCV Plus Unity*. Accessed: Apr. 30, 2025. [Online]. Available: https://assetstore.unity.com/packages/tools/integration/opencv-plus-unity-85928

[36] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981, doi: 10.1145/358669.358692.

[37] J. Brooke, "SUS: A 'Quick and Dirty' usability scale," in *Usability Evaluation in Industry*. U.K.: Taylor & Francis, 1996, p. 189.

[38] C. P. Walmsley, S. A. Williams, T. Grisbrook, C. Elliott, C. Imms, and A. Campbell, "Measurement of upper limb range of motion using wearable sensors: A systematic review," *Sports Med. Open*, vol. 4, no. 1, p. 53, Dec. 2018, doi: 10.1186/s40798-018-0167-7.

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - ROR identifier: 00x0ma614

2.2.2 Key Findings and Implications

The reproduced article demonstrates that accurate 3D elbow angle estimation can be achieved using a low-cost, vision-based approach without reliance on depth sensors, wearable devices, or complex AI models. The system achieved clinically acceptable accuracy within relevant ranges of motion and maintained robustness under different calibration and environmental conditions.

A key contribution of this work is the use of a geometric, marker-based tracking strategy, which offers advantages in terms of computational efficiency, interpretability, and robustness to occlusion when compared to AI-based pose estimation methods. The comparative analysis with MediaPipe reinforces the suitability of this approach for rehabilitation scenarios involving challenging arm positions.

The study also identifies important practical factors affecting performance, such as camera distance, marker size, spatial positioning within the image, and movement velocity. These findings provide concrete guidelines for system configuration and contribute to the reliability of real-world deployment.

In conclusion, this work addresses one of the main limitations identified in literature, the dependence on specialized hardware, by demonstrating that clinically relevant kinematic data can be captured using accessible technologies. This establishes a technological foundation for the next phase of this thesis, which focuses on translating such kinematic measurements into automated clinical assessment outcomes.

2.3 Preliminary validation study

2.3.1 Context and Summary

This section presents the full reproduction of the article entitled *AI-driven low-cost rehabilitation exergame as a lightweight framework for stroke assessment*, published in npj Digital Medicine (TANNÚS; VALENTINI; NAVES, 2026). This article is distributed under a Creative Commons Attribution 4.0 International License, which permits reproduction, distribution, and adaptation, provided appropriate credit is given to the original authors.

This article corresponds to Phase 3 of the research developed in this doctoral thesis and represents the full development of the proposed framework. While the previous phase focused on accurate motion capture using a low-cost vision-based system, this study advances the approach by integrating automated clinical assessment directly into the rehabilitation exergame, eliminating the need for separate evaluation procedures.

The main objective of this work is to investigate whether interpretable kinematic features extracted during gameplay can reliably estimate clinical motor scores, specifically the FMA, which is a gold standard in post-stroke motor evaluation but is time-consuming and requires trained clinicians.

To achieve this, the proposed system uses a standard RGB camera and the MediaPipe framework to extract 2D hand and arm trajectories during gameplay. From these data, 16 kinematic and spatiotemporal features were computed, including hand angle, range of motion, movement area, traveled distance, and shoulder-elbow coordination. These features were designed to be clinically interpretable and computationally lightweight.

It is important to note that, although the previous phase demonstrated that a marker-based geometric approach can achieve higher elbow angle accuracy and robustness in controlled scenarios, the adoption of an AI-based tracking method in this study was a decision to completely eliminate external hardware, such as markers or calibration objects, thereby simplifying system setup and improving scalability, usability, and accessibility in real-world scenarios.

The study involved 12 post-stroke participants (24 upper limbs), with motor impairment levels stratified according to FMA scores. Several variables, such as average hand angle, hand range of motion, coordination, and spatial exploration, demonstrated clear separation across severity groups, indicating their potential as digital biomarkers.

Correlation analyses revealed that multiple gameplay-derived features were significantly associated with FMA scores, with some variables (e.g., hand angle and joint coordination) showing strong monotonic relationships with motor function. These findings support the hypothesis that movement patterns captured during gameplay reflect clinically meaningful aspects of motor impairment.

To estimate clinical scores, multiple linear regression models were developed using combinations of selected features. The best-performing model achieved high predictive performance (Spearman $\rho = 0.92$, $R^2 = 0.89$, RMSE = 4.42), indicating strong agreement with clinical evaluations. Additionally, the model was able to classify patients into severity levels (mild, moderate, severe) with accuracies ranging from 86% to 93%.

The study also preliminarily compared the proposed approach with more complex machine learning models. Despite their complexity, these models did not outperform the lightweight regression approach, probably due to low sample data, therefore, no conclusions could be made.

Overall, this article demonstrate that automated motor assessment can potentially be embedded within a low-cost rehabilitation exergame, providing clinically meaningful evaluation in real time and without external sensors.

npj | digital medicine

Article

Published in partnership with Seoul National University Bundang Hospital

https://doi.org/10.1038/s41746-026-02383-1

AI-driven low-cost rehabilitation exergame as a lightweight framework for stroke assessment

Check for updates

Júlia Tannús ✉, Caroline Valentini & Eduardo Naves

Stroke is a leading cause of long-term disability, often affecting upper-limb motor function and requiring continuous assessment. The Fugl-Meyer Assessment (FMA), though a clinical gold standard, is time-consuming and demands specialized personnel. This study presents an AI-driven, low-cost rehabilitation exergame that simultaneously provides therapy and automatically estimates upper-limb motor performance during gameplay using only a standard camera. Sixteen kinematic and spatiotemporal features were extracted from 2D hand and arm trajectories of twelve post-stroke individuals (24 limbs, 14 affected) using the MediaPipe framework. Features such as hand angle, range of motion, movement area, traveled distance, and shoulder–elbow coordination showed strong correlations with FMA scores and stratified participants by motor severity. A lightweight linear regression model achieved high predictive performance (Spearman $\rho = 0.92$, $R^2 = 0.89$, RMSE = 4.42) and classified severity levels with 86–93% accuracy. This interpretable approach outperformed complex machine learning models, highlighting the clinical relevance of transparent metrics embedded in gameplay. The proposed framework is sensor-free, scalable, and reproducible, offering immediate feedback while reducing clinical workload and enabling accessible digital biomarkers for telerehabilitation and remote monitoring after stroke.

Stroke remains a leading cause of long-term disability worldwide, with over 12 million new cases annually and approximately 100 million stroke survivors globally, many of whom experience upper-limb motor deficits that hinder daily function[1,2].

For rehabilitation, while in-person treatments remain essential, they can be time-consuming, resource-intensive, and limited in availability[3]. These programs can also be tedious and financially burdensome, requiring patient transportation to clinical sites. In contrast, video game-based rehabilitation has emerged as a motivational and accessible alternative[4,5]. Virtual reality games have proven capable of engaging stroke patients in repetitive motor tasks that support neuroplasticity and motor recovery[6–9].

In parallel, accurate assessment of motor impairment is essential for treatment planning and recovery monitoring. The Fugl-Meyer Assessment (FMA) is widely considered a gold standard for quantifying post-stroke motor function[10–12], but it is time-consuming, subjective, and requires trained clinicians[3,13,14].

For these reasons, there are some initiatives to automate motor assessments, including smartphone apps[15,16], motion sensor-based systems[17,18], and machine learning models that estimate motor scores from movement features[19–21].

While many of these systems are promising, they often require dedicated assessment time, lack interpretability, need external sensors (such as depth cameras or inertial sensors), or rely on complex architectures such as 3D pose estimation or deep neural networks[5,22]. In contrast, combining motor assessment directly into the rehabilitation game itself, as in the present work, offers a compelling advantage: it eliminates the need for separate clinical evaluations, reduces monotony, increases engagement, and enables real-time, high-frequency tracking of recovery without burdening therapists or patients. This integration promotes scalability and personalization, allowing stroke survivors to be continuously monitored during gameplay using simple, transparent metrics that reflect functional performance.

Unlike conventional assessments that must be administered separately, the proposed system performs evaluation during therapeutic gameplay itself. As the patient engages in motor exercises, the same movement data used for rehabilitation are automatically analyzed to generate FMA-equivalent digital scores, eliminating the need for additional assessment time.

In this framework, the term "AI-driven" refers to the use of Google's MediaPipe: a computer-vision framework powered by machine learning for real-time hand and body tracking. By leveraging this AI-based system, the

Assistive Technologies Group, Federal University of Uberlândia, Uberlândia, Brazil. ✉e-mail: julia.tannus@ufu.br

https://doi.org/10.1038/s41746-026-02383-1

same standard camera used for gameplay also performs motion capture, removing the dependence on specialized hardware. The resulting kinematic features serve as AI-driven digital biomarkers, offering objective and transparent indicators of motor function while maintaining a low-cost and scalable setup.

Therefore, this work proposes a lightweight, AI-based exergame system for post-stroke rehabilitation that also estimates upper-limb motor function using only simple, interpretable kinematic features extracted from 2D wrist and hand movements during gameplay. In this paper, it was tested how well these features reflect clinical severity, discriminate across FMA strata, and predict motor scores. By leveraging basic movement metrics in a linear equation, such as hand aperture, spatial exploration, and joint coordination, without relying on expensive hardware or black-box deep learning models, this system offers a cost-effective and accessible alternative to traditional assessments.

Table 1 | Details of the stroke subjects

Subject	Age (years)	Gender	Affected Side	FMA[a] Left	FMA[a] Right
S1	66	Male	Left	23	62
S2	36	Male	Left	15	66
S3	33	Female	Right	66	62
S4	53	Male	Both	48	29
S5	45	Male	Both	26	17
S6	21	Male	Right	66	33
S7	46	Male	Left	43	66
S8	49	Female	Left	22	66
S9	61	Female	Left	21	66
S10	37	Male	Left	48	66
S11	52	Male	Left	13	66
S12	61	Male	Right	66	15

[a]FMA refers to the Fugl-Meyer Assessment Upper Extremity, a score between 0 (no function) and 66 (intact).

Results

Participant characteristics

Twelve individuals with chronic stroke participated in the study, including nine males and three females. Participants ranged in age from 21 to 66 years, with a mean age of 46.7 years (SD = 13.2). The time since stroke ranged from four months to 30 years. Most participants had experienced ischemic stroke ($n = 10$), while two had suffered hemorrhagic stroke. In two cases, bilateral upper-limb paresis was reported.

Each participant performed gameplay sessions using both upper limbs and was assessed using the FMA on each side. This resulted in a total of 24 limb-level observations, enabling intra-individual comparisons. Table 1 presents the demographic and clinical characteristics of the participants, including age, gender, affected side, and FMA scores for each limb.

Based on FMA scores, the tested limbs were stratified into four categories: (i) severe impairment (FMA ≤ 20), (ii) moderate impairment (21 ≤ FMA ≤ 45), (iii) mild impairment (FMA > 45), and (iv) control limbs (non-affected side of participants with unilateral hemiparesis). Table 2 summarizes the distribution of upper limbs across these severity levels, including the number of limbs per group, average age, and mean FMA scores.

Feature summary

Gameplay-derived features were standardized (Z-scores) and compared across four severity groups (Severe, Moderate, Mild, Control). Figure 1

Table 2 | Upper limbs tested stratified by severity level

Level of Hemiparesis	Number of Limbs	Left Side	Average Age (years)	Average FMA Score
Severe	4	2	48.5 ± 10.6	15.0 ± 1.6
Moderate	7	5	50.3 ± 7.4	28.0 ± 10.2
Mild	3	2	41.0 ± 10.6	52.7 ± 8.1
Control	10	3	39.0 ± 19.1	65.0 ± 2.0

Each participant contributed two upper limbs to the dataset (one left and one right), for a total of 24 limbs. Limbs were classified according to their motor function based on the FMA: (i) Severe: FMA ≤ 20, (ii) Moderate: FMA 21-45, (iii) Mild: FMA > 45 (affected limb), and (iv) Control: non-affected limb.

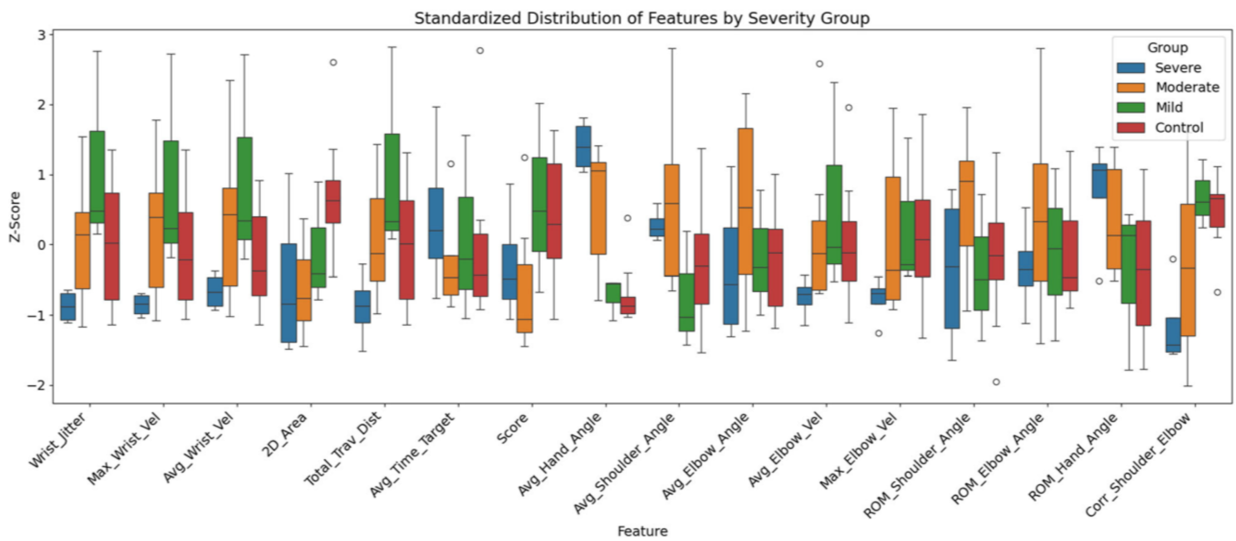


Fig. 1 | Standardized distributions of gameplay-derived features stratified by upper-limb severity level. The plot shows Z-scores for all 16 kinematic features and the in-game score, grouped by severity. Features such as Avg_Hand_Angle, ROM_Hand_Angle, Corr_Shoulder_Elbow, and 2D_Area demonstrate clear separation between groups, with progressively higher median values from Severe to Control. Boxes represent interquartile ranges (IQR), with whiskers indicating 1.5×IQR and dots representing outliers. For each feature, the degree of separation between the boxes indicates how well that variable discriminates among severity categories: the less overlap, the better its classification capability. The direction of the median values (whether they increase or decrease across groups) shows whether the relationship with motor function is linear and positive or negative, that is, whether higher values represent better or worse performance.

https://doi.org/10.1038/s41746-026-02383-1

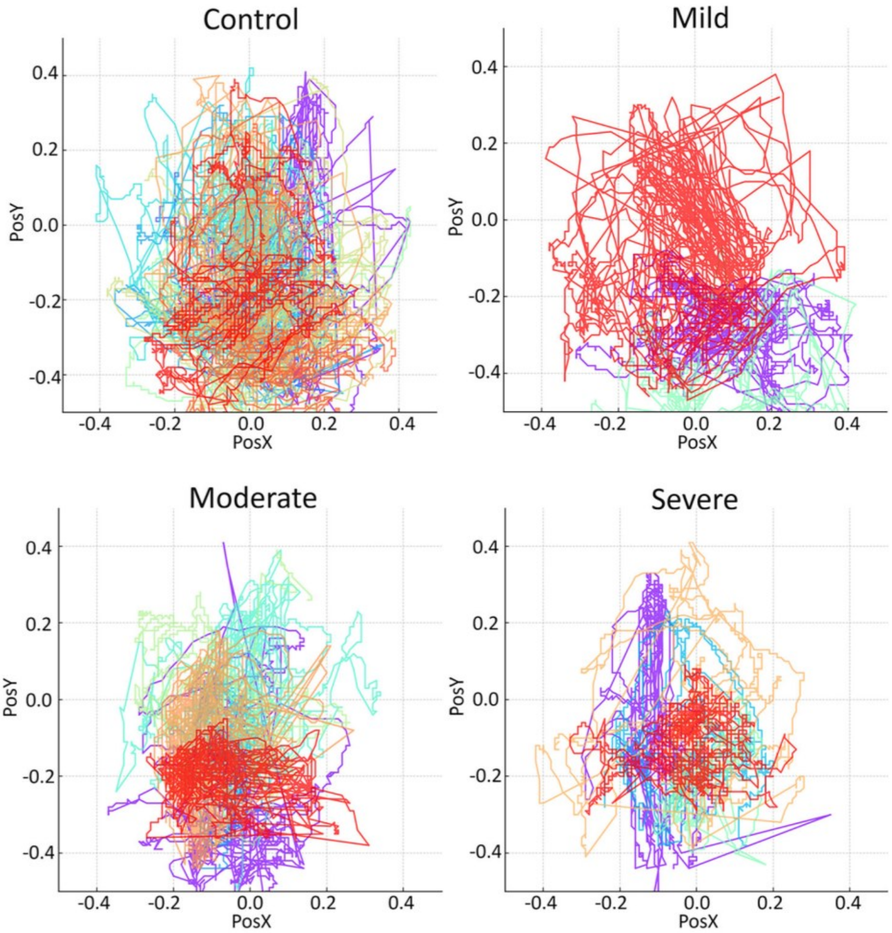


Fig. 2 | Frame-by-frame wrist trajectories captured during gameplay, stratified by motor severity level based on Fugl-Meyer scores. Limbs in the Control group explored a significantly larger portion of the 2D workspace, demonstrating well-distributed trajectories and full utilization of the screen area, supporting the discriminative power of the *2D_Area* feature.

presents a boxplot summarizing the distribution of all 16 features across these categories.

Several features demonstrated consistent and discriminative trends. Specifically, *Avg_Hand_Angle*, *ROM_Hand_Angle*, *Corr_Shoulder_Elbow*, and *2D_Area* exhibited progressive and monotonic changes in median values across severity levels, with *Avg_Hand_Angle* standing out as the most clearly stratified variable.

Conversely, features such as *Avg_Wrist_Vel*, *ROM_Shoulder_Angle*, and *Avg_Elbow_Angle* showed minimal group separation and high within-group dispersion, indicating limited discriminative power in isolation.

Overall, the feature distribution analysis reveals three key conclusions:

1. Certain features consistently reflect clinical severity, particularly *Avg_Hand_Angle*, *ROM_Hand_Angle*, *Corr_Shoulder_Elbow*, and *2D_Area*, making them strong candidates for predictive modeling and digital assessment tools.
2. Some features, despite high variability, still provide partial stratification, suggesting they may be useful when combined into multivariate models.
3. Some metrics are uninformative and should be used cautiously.

Trajectory plots (Fig. 2) showed compact, centralized movements in the Severe group versus broader exploration in Controls, with Mild and Moderate groups presenting intermediate patterns. The results reinforce *2D_Area* as a clinically interpretable indicator of upper-limb functional reach in post-stroke individuals.

Correlation with clinical score per feature

To explore how gameplay-derived features relate to motor function, Spearman correlations were computed between each feature and the FMA score, stratified by limb type (affected, control), and considering all limbs together (total). The results are presented in Fig. 3.

The following features showed statistically significant correlation with FMA, ordered by the absolute value of Spearman's ρ. Significant correlations are highlighted using: *($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$):

1. Avg_Hand_Angle (total): $\rho = -0.76$ ***
2. Corr_Shoulder_Elbow (affected): $\rho = 0.67$ **
3. Avg_Hand_Angle (affected): $\rho = -0.61$ *
4. Corr_Shoulder_Elbow (total): $\rho = 0.59$ **
5. Total_Trav_Dist (affected): $\rho = 0.54$ *
6. 2D_Area (total): $\rho = 0.50$ *
7. Score (total): $\rho = 0.49$ *
8. ROM_Hand_Angle (total): $\rho = -0.47$ *

These results confirm that features related to joint coordination, hand orientation, trajectory coverage, and game performance are consistently associated with clinical motor scores. In particular, the Average Hand Angle and Shoulder-Elbow Coordination show robust associations across both affected limbs and the entire dataset.

Predictive modeling

To explore the predictive capacity of the extracted features in estimating clinical motor function, multiple linear regression analyses were performed using exhaustive feature selection with all 16 features, with maximum of 5 features per equation, to avoid overfitting. This approach systematically evaluates all possible combinations of features to identify the subset that maximizes correlation with the ground truth FMA scores.

Separate models were constructed for the affected limb and for the total group. For each group, the best regression models were identified according

https://doi.org/10.1038/s41746-026-02383-1

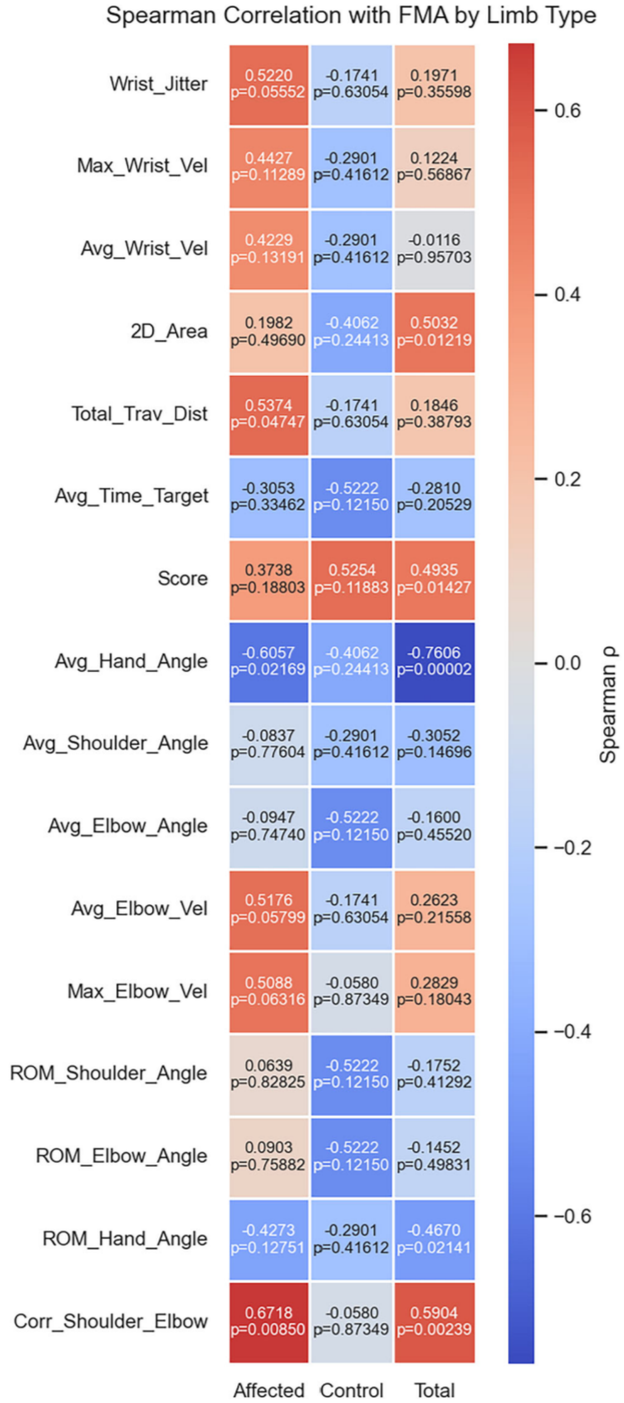


Fig. 3 | Spearman correlation coefficients (ρ) and corresponding *p*-values between kinematic features and FMA scores, stratified by limb type (Affected, Control, and Total). Cells are colored according to the magnitude and direction of ρ, with warmer tones indicating positive correlations and cooler tones indicating negative correlations. For example, a dark blue in the Avg_Hand_Angle group - total and a very warm red in the Corr_Shoulder_Elbow group - affected indicate the characteristics with the highest correlation.

to Spearman's ρ, RMSE, and R^2. All models were derived from standardized (Z-scored) features. The top-performing equations, along with their corresponding performance values, are summarized in Table 3.

For the affected limb group, the optimal multiple regression models varied depending on the evaluation metric. Using a 50/50 hold-out split, the model maximizing Spearman correlation achieved an excellent association with the clinical score ($\rho = 0.99$, $p = 0.0003$), although with moderate explained variance ($R^2 = 0.52$) and RMSE = 9.89. By contrast, the model optimized for error reduction reached the best overall fit ($R^2 = 0.89$, RMSE = 4.42), still maintaining a strong correlation ($\rho = 0.92$, $p = 0.003$). LOOCV (leave-one-out cross validation) models for the affected subset produced lower correlations ($\rho = 0.67$–0.82) and explained variance ($R^2 = 0.32$–0.55), confirming that the hold-out strategy yielded more stable and clinically relevant predictions in this small sample.

For the total group, results were more consistent. The hold-out models performed robustly, with the best solution balancing high correlation ($\rho = 0.94$, $p < 0.001$) and strong explained variance ($R^2 = 0.83$) while maintaining an acceptable error (RMSE = 8.92). LOOCV models again performed slightly worse, particularly in terms of error, reinforcing the advantage of hold-out validation for this dataset.

Importantly, the features selected through exhaustive feature selection did not always correspond to those with the strongest individual correlations with the FMA. This highlights that exhaustive feature selection identified not only individually strong predictors but also synergistic feature sets that together maximized predictive performance. Accordingly, variations in coefficients across equations arise from the inclusion of different predictor combinations identified through exhaustive feature selection. These alternative formulations highlight multiple relevant feature sets rather than instability, reflecting the exploratory nature of this proof-of-concept study. Thus, while univariate correlations provided useful guidance, the multivariate regression models offered a more comprehensive view of how gameplay-derived metrics map onto clinical motor outcomes.

The final regression model for the affected group, summarized in Table 3 and illustrated in Fig. 4, achieved $R^2 = 0.89$ and Spearman's $\rho = 0.92$, representing the best overall balance among the three performance metrics (R^2, RMSE, and Spearman correlation). Earlier exploratory models with higher ρ but lower R^2 were excluded from the final results to ensure consistency between rank and variance-based measures.

Exploratory diagnostic accuracy

The confusion matrix in Table 4 demonstrates strong classification performance across severity levels. Moderate cases were most consistently recognized, with all 7 instances correctly classified. Mild cases showed agreement in 2 out of 3 participants, while severe cases were correctly identified in 3 out of 4 participants. Importantly, misclassifications occurred only between adjacent severity levels (e.g., mild vs. moderate, moderate vs. severe), with no extreme errors across non-neighboring classes.

As summarized in Table 5, the regression equation achieved balanced performance across all groups. The Mild class obtained perfect precision (1.00) but slightly lower recall (0.67), resulting in an F1-score of 0.80 and overall accuracy of 0.93. The Moderate class showed the highest recall (1.00) with good precision (0.78), achieving an F1-score of 0.88 and accuracy of 0.86. The Severe class also reached high performance, with precision of 1.00, recall of 0.75, and an F1-score of 0.86, corresponding to an accuracy of 0.93. These results indicate that the regression-based model generalized well across severity categories, reliably distinguishing between mild, moderate, and severe impairment levels.

Machine learning analysis

The Random Forest model demonstrated robust predictive capacity when trained on the total dataset, achieving $\rho = 0.76$ ($p = 0.030$), $R^2 = 0.63$, and RMSE = 12.25 FMA points in hold-out validation.

The temporal deep learning model (CNN1D+BiLSTM) achieved modest performance on the independent arm-level test set ($R^2 = 0.59$, RMSE = 11.5, $\rho = 0.75$; $p = 0.05$), underscoring the challenge of generalization with small datasets despite promising validation results.

Regularized linear models yielded highly interpretable sparse solutions. Both Lasso ($\alpha = 7.05$) and Elastic Net converged to the same two predictors: *Score* (positive coefficient) and *Avg_Hand_Angle* (negative coefficient), and achieved comparable performance ($R^2 = 0.53$, RMSE = 13.7, $\rho = 0.83$;

Table 3 | Best multiple linear regression models by limb group, validation method and best metric (z-score normalized)

Group	Validation Method	Best Metric	Equation	Spearman ρ (p)	RMSE	R^2
Affected	Hold-out (50/50)	Spearman ρ	FMA = -14.57·Avg_Hand_Angle + 11.10·Corr_Shoulder_Elbow + 8.33·Avg_Time_Target + -7.14·Avg_Elbow_Angle + -10.10·Max_Elbow_Vel + 30.58	0.99 (0.0003)	9.89	0.52
Affected	Hold-out (50/50)	RMSE, R^2	FMA = -11.26·Avg_Hand_Angle + 16.32·Total_Trav_Dist + -13.57·Wrist_Jitter + -1.77·ROM_Shoulder_Angle + 29.64	0.92 (0.003)	4.42	0.89
Affected	LOOCV	Spearman ρ	FMA = -7.08·Avg_Hand_Angle + 10.44·Corr_Shoulder_Elbow + -5.32·2D_Area + 1.44·Avg_Time_Target + 30.58	0.82 (0.001)	12.78	0.32
Affected	LOOCV	RMSE	FMA = 13.79·Corr_Shoulder_Elbow + -5.38·2D_Area + -14.56·Avg_Shoulder_Angle + -8.56·Avg_Elbow_Angle + 11.71·ROM_Shoulder_Angle + 29.64	0.67 (0.008)	10.04	0.52
Affected	LOOCV	R^2	FMA = -14.57·Avg_Hand_Angle + 11.10·Corr_Shoulder_Elbow + 8.33·Avg_Time_Target + -7.14·Avg_Elbow_Angle + -10.10·Max_Elbow_Vel + 30.58	0.72 (0.008)	10.38	0.55
Total	Hold-out (70/30)	Spearman ρ	FMA = 12.53·2D_Area + 7.61·Score + -6.42·ROM_Shoulder_Angle + 44.62	0.94 (0.0005)	14.12	0.42
Total	Hold-out (70/30)	RMSE, R^2	FMA = -14.82·Avg_Hand_Angle + 5.85·Corr_Shoulder_Elbow + 12.51·Score + -7.63·Avg_Wrist_Vel + 8.13·Avg_Time_Target + 46.50	0.93 (0.003)	8.92	0.83
Total	LOOCV	Spearman ρ	FMA = -15.77·Avg_Hand_Angle + 2.98·2D_Area + 18.65·Total_Trav_Dist + -19.99·Avg_Wrist_Vel + 0.74·Avg_Time_Target + 46.50	0.87 (0.000009)	12.88	0.62
Total	LOOCV	RMSE, R^2	FMA = -15.45·Avg_Hand_Angle + 4.67·2D_Area + 13.77·Score + -6.74·Avg_Wrist_Vel + 7.92·Avg_Time_Target + 46.50	0.80 (0.000009)	10.67	0.74

LOOCV Leave-one-out cross validation. The best overall models for the Affected (*n* = 14) and Total (*n* = 24) group are highlighted in green.

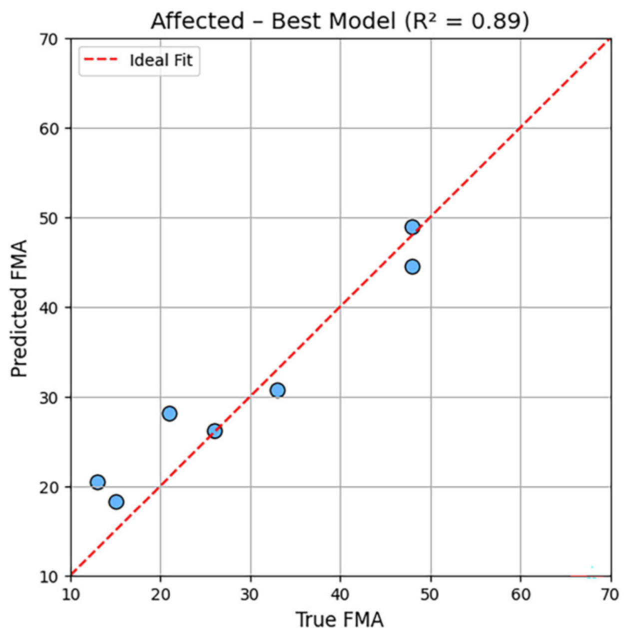


Fig. 4 | Predicted FMA values from the best multiple linear regression model for the affected limb subset. The red dashed line represents the ideal fit (y = x).

p = 0.011). Although they explained slightly less variance than Random Forest, their strong monotonic association with FMA and parsimony make them particularly appealing for clinical translation. Summarized results are in Table 6. Importantly, the best result did not surpass the best linear regression models.

This evaluation was done in the total group, because the affected group did not wield statistically significant results, due to low sample size, which also highlights the regression model superiority.

Table 4 | Confusion matrix by class (affected)

	Predicted				
		Mild	Moderate	Severe	Total
True	Mild	2	1	0	3
	Moderate	0	7	0	7
	Severe	0	1	3	4
	Total	2	9	3	14

Table 5 | Best regression equation performance by class (affected)

Class	Precision	Recall	F1-Score	Accuracy
Mild	1.00	0.67	0.80	0.93
Moderate	0.78	1.00	0.88	0.86
Severe	1.00	0.75	0.86	0.93

Table 6 | Summary of machine learning models applied (total group)

Model	R^2	RMSE	Spearman ρ
Random Forest	0.63	12.25	0.76 (*p* = 0.030)
CNN1D+BiLSTM	0.59	11.5	0.75 (*p* = 0.05)
Lasso (α ≈ 7.05)	0.54	13.7	0.83 (*p* = 0.011)
Elastic Net	0.53	13.8	0.83 (*p* = 0.011)

Machine learning models were included as exploratory analyses to provide contextual comparison, and not as the primary or final methodological objective of the study. They were not prioritized due to their limited feasibility for real-time deployment within a mobile exergame environment.

https://doi.org/10.1038/s41746-026-02383-1

Discussion

This study evaluated the feasibility and clinical relevance of an AI-driven, low-cost rehabilitation exergame for assessing upper-limb motor function in individuals with chronic stroke. The results demonstrated that gameplay-derived kinematic features not only reflect clinically interpretable differences across severity levels but also predict the gold-standard FMA scores with high accuracy, highlighting the system's potential as a transparent and scalable digital biomarker.

The exergame simultaneously serves therapeutic and evaluative purposes, capturing clinically interpretable kinematic data during gameplay without requiring separate assessment sessions. This dual function enhances efficiency and allows objective, high-frequency monitoring of recovery progress.

Identifying features that correlate with FMA is essential for building interpretable and clinically interpretable assessment tools. In this study, several metrics stood out for their discriminative power across severity levels.

Average hand angle and hand angle range of motion were particularly informative, capturing the ability to open the hand, a key clinical deficit in severe cases, often associated with clenched hands or spasticity. Their progressive changes across severity groups reinforce hand aperture as a central marker of motor impairment. In the regression equations, the negative coefficient of *Avg_Hand_Angle* indicates that increased finger flexion (clenched-hand posture) is associated with lower FMA scores, consistent with clinical patterns of spasticity. Conversely, positive coefficients such as those for *Corr_Shoulder_Elbow* and *2D_Area* denote more coordinated and spatially extensive movements, which align with better voluntary control and higher functional scores. These relationships strengthen the interpretability of the model and its translational relevance for clinical assessment.

Spatial exploration features, such as 2D movement area and total traveled distance, also showed strong associations with FMA, reflecting how higher-functioning participants performed broader, more frequent movements. Shoulder-elbow coordination further distinguished between synergy-dominated patterns in low-FMA patients and refined joint control in those with better outcomes.

Interestingly, features traditionally used in motor assessment, such as velocity or raw joint angles, were not among the best correlated features. This may be due to clinical heterogeneity: patients may be spastic or flaccid, limiting the generalizability of angular metrics. Velocity, in particular, proved ambiguous, since high values may reflect either controlled speed in higher-functioning participants or poorly coordinated, abrupt movements in those with impairments. As a result, speed alone did not reliably indicate motor capacity. These findings indicate that features that capture core motor mechanisms can be more valuable than less interpretable metrics.

Multiple linear regression models built on the most relevant features achieved highly accurate predictions of FMA scores, particularly in the affected limb group ($\rho = 0.92$, RSME = 4.42, $R^2 = 0.89$). Even with hold-out validation, the models retained strong predictive power, confirming that motor function can be reliably inferred from simple combinations of movement metrics.

Notably, more complex machine learning approaches, including Random Forest and temporal deep learning models (CNN1D+BiLSTM), did not surpass the performance of the lightweight linear regression. These advanced models showed promising internal validation but generalized poorly on independent test sets, underscoring the challenge of applying black-box architectures to small datasets. Regularized linear models such as Lasso and Elastic Net offered interpretability but achieved lower accuracy than the proposed regression approach.

This finding highlights a central innovation of the framework: a simple, interpretable functional evaluation algorithm, directly integrated into post-stroke rehabilitation gameplay, can outperform more complex machine learning pipelines, making the system more reliable, clinically transparent, and easier to adopt in real-world. This design ensures low computational cost, allowing on-device analysis and continuous patient monitoring during rehabilitation sessions. The ability to assess motor function as the patient plays removes the need for standalone assessment software and supports more frequent, naturalistic evaluations.

The small cohort size limited the ability to distinguish performance differences among machine learning models. However, this constraint also highlights a practical advantage of the proposed lightweight regression model, which achieved strong correlations and clinically relevant predictions with relatively few training samples.

A larger, follow-up study is planned to expand the dataset and validate these exploratory findings, enabling robust benchmarking across models trained with homogeneous data representations.

To contextualize our findings, Table 7 summarizes representative studies that used different sensing modalities and modeling approaches for automated upper-limb motor assessment. Most previous systems rely on specialized hardware such as depth cameras, inertial sensors, or multi-sensor fusion, and often apply black-box architectures that hinder interpretability. In contrast, the present framework achieves comparable correlations with clinical scores using only the standard RGB camera from the display device (e.g., laptop, tablet) and a transparent regression model embedded directly within gameplay.

This combination of AI-based motion capture and lightweight, interpretable evaluation demonstrates how accessible, low-cost solutions can provide clinically interpretable digital biomarkers for telerehabilitation.

The exploratory diagnostic accuracy analysis further demonstrated that predicted FMA values could stratify patients into clinical severity groups with 86–93% accuracy, misclassifying only between neighboring categories. This level of agreement is clinically acceptable and reinforces the framework's potential to support automated functional classification.

Unlike previous systems requiring Kinect, Wii, depth cameras, inertial sensors, or exoskeletons, the proposed exergame achieves assessment using only a standard camera. By embedding evaluation into gameplay itself, the system eliminates the need for additional software or separate testing sessions. This sensor-free, low-cost, and scalable approach reduces setup complexity, enhances patient engagement, and lowers clinical workload, paving the way for widespread telerehabilitation applications.

This study introduced a lightweight, AI-driven rehabilitation exergame capable of simultaneously engaging patients in therapy and assessing upper-limb motor function. Specific gameplay-derived features, particularly hand aperture, 2D movement area, and shoulder-elbow coordination, showed strong associations with the FMA, enabling accurate prediction of clinical scores and stratification of motor severity.

Crucially, a simple linear regression model provided superior performance and interpretability compared to more complex machine learning approaches, demonstrating that transparency and clinical usability can outweigh the marginal gains of black-box algorithms in small-scale rehabilitation studies.

Table 7 | Comparison of related studies (fma-ue 33 items)

Study	Sensor Type	Modeling Approach	Dataset Size	Main Result
Song et al.[26]	Camera and inertial sensors of a smartphone	Decision trees	$n = 10$	$R^2 = 0.78$
Chen et al.[20]	Custom optical capture device	Decision trees	$n = 79$	RMSE = 17.4
Jiang et al.[27]	Microsoft Kinect	Fuzzy inference	$n = 25$	Accuracy = 93.5%
Present study	Camera of the game display device (MediaPipe) – no external sensors	Linear regression	$n = 24$	$R^2 = 0.89$

https://doi.org/10.1038/s41746-026-02383-1

The framework's low-cost and sensor-free design is a major innovation: by relying solely on a standard camera, it avoids the expense and logistical burden of external sensors, such as Kinect, inertial units, or exoskeletons. Furthermore, because assessment occurs in real time during gameplay, the system eliminates the need for dedicated evaluation sessions or post-processing software, offering immediate feedback and reducing clinician workload.

Taken together, these features position the proposed exergame as a practical, scalable, and innovative tool for stroke rehabilitation and remote monitoring. While the small sample size and cross-sectional design limit generalizability, this proof-of-concept study demonstrates strong potential for integration into telerehabilitation settings. Future work should validate the framework in larger and more diverse populations, explore longitudinal responsiveness to recovery, and investigate its role in personalized rehabilitation pathways. Also, multiple independent raters and inter-rater reliability analyses will be included to enhance robustness.

Methods

Exergame system

The exergame employed in this study was developed using the Unity 2020.1.17 game engine (https://unity.com/) and offers a dynamic and engaging environment aimed at evaluating upper-limb motor function. The gameplay centers around guiding a bird avatar to collect fruits dispersed across a virtual landscape (Fig. 5). The game was created to entertain users and enable rehabilitation, not to resemble a clinical evaluation.

Player control is based on real-time 3D hand tracking, implemented through the MediaPipe framework[23], which operates with a standard camera, commonly found on devices such as tablets, laptops, and smartphones.

MediaPipe relies on Artificial Intelligence-based models for real-time hand and body landmark detection, using deep neural networks trained on large-scale human pose datasets. This AI-driven architecture enables accurate 3D tracking directly from standard RGB images, allowing the same low-cost camera used for gameplay to perform motion capture. By eliminating the need for depth cameras, inertial sensors, or external hardware, MediaPipe provides an accessible, scalable, and cost-effective foundation for automated rehabilitation assessment.

Specifically, the X and Y coordinates (in pixels) of the wrist landmark are mapped to the character's horizontal and vertical movement axes within the virtual environment (Fig. 6).

The bird's velocity increases proportionally to the wrist's distance from the screen center, thereby stimulating participants to explore a larger range of motion.

This two-dimensional control strategy was selected to resemble functional components commonly observed in clinical motor evaluations, particularly those involving planar reaching and positioning tasks, such as lifting the arm toward a target or bringing the hand to the mouth. This design not only facilitates the interpretation of gameplay-derived metrics in relation to clinical scores, but also ensures greater accessibility for individuals with reduced motor capacity or fatigue.

A key advantage of this tracking method is the elimination of wearable sensors or physical markers, which simplifies the setup and can reduce implementation costs.

Participants were instructed to maintain the hand extended and parallel to the camera, which allowed the system to extract features such as mean finger extension and range of motion, while ensuring hand visibility for tracking.

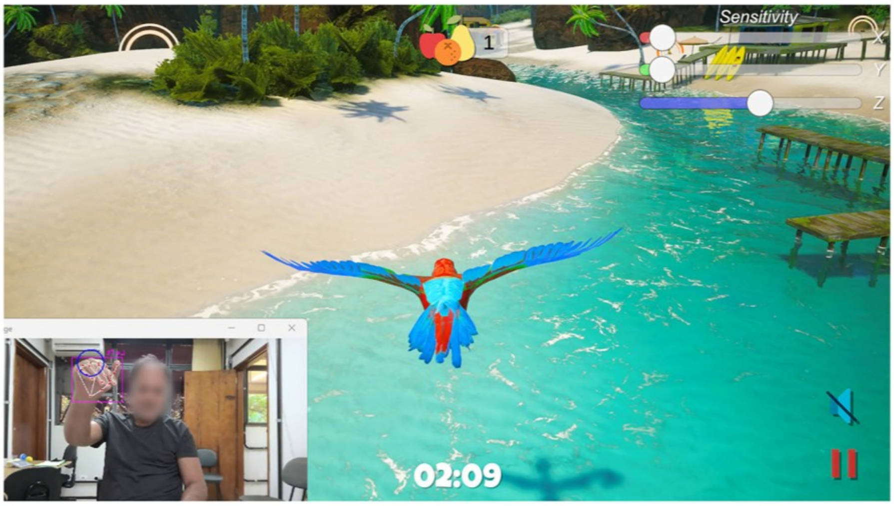


Fig. 5 | Screenshot of the exergame during gameplay. The main scene displays the bird avatar navigating a tropical environment while collecting virtual fruits. The participant's wrist position in the X and Y axes (in pixels), captured in real-time via MediaPipe hand tracking, is used to control the character's movement. Source: The authors. Written consent to publish the image was obtained from the participant.

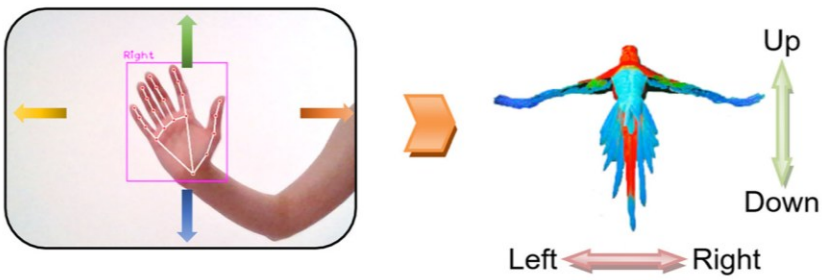


Fig. 6 | Representation of the hand tracking mechanism used in the exergame system. The MediaPipe framework detects 21 hand landmarks, from which the wrist X and Y position (landmark 0), in pixel units, is extracted to control the avatar's movement, enabling navigation within the virtual environment. Source: The authors. Edited using Microsoft Word.

https://doi.org/10.1038/s41746-026-02383-1

Fig. 7 | Administration of the FMA. A trained physical therapist evaluates motor function in a post-stroke participant through standardized tasks involving the shoulder, elbow, forearm, wrist, and hand. The assessment was conducted bilaterally in a controlled clinical setting. Source: The authors. Written consent to publish the image was obtained from the participant.

Clinical scale

The FMA[24] is a well-established clinical instrument commonly employed to assess motor deficits in individuals recovering from stroke. It comprises a set of items that examine reflexes as well as isolated and synergistic movements across various segments of the upper limb, including the shoulder, elbow, forearm, wrist, and hand. The total score ranges from 0 to 66, with higher values indicating superior motor performance. In this study, the FMA was applied bilaterally, on both the impaired and non-impaired upper limbs, to enable intra-individual comparisons and to evaluate asymmetries in motor function (Fig. 7). All FMA assessments were conducted by a single licensed physical therapist with more than ten years of experience in neurorehabilitation. The evaluator followed standardized scoring procedures and was blinded to the gameplay and data processing results to minimize bias. Written informed consent for publication of the images was obtained from all participants.

Data collection

All evaluations were performed during a single experimental session to ensure consistent testing conditions, minimize fatigue or learning effects, and avoid external variables that could compromise data reliability. To ensure that participants were assessed in their baseline motor conditions, the FMA was administered before the exergame task.

To minimize potential bias in the statistical analyses, the application of the FMA was blinded to the individual responsible for data processing and feature extraction.

Each eligible participant completed three minutes of two consecutive gameplay sessions, one using the affected upper limb and another using the unaffected limb.

During gameplay, the system captured data continuously and frame by frame, as follows:

- The 3D coordinates (X, Y, Z) of the 21 hand landmarks detected by the MediaPipe Hands framework;
- The shoulder and elbow joint angles, estimated from the MediaPipe Pose landmarks;
- The game score, defined by the number of fruits successfully collected by the avatar;
- The elapsed time since the beginning of the session, in seconds.

All raw gameplay data were exported in .csv format and analyzed using custom Python scripts to extract key kinematic metrics, such as range of motion, movement smoothness, and average velocity.

Participants

Participants were recruited from the clinical referral network of the Assistive Technology Laboratory at the Federal University of Uberlândia, Brazil, forming a convenience sample. The sample size was selected based on general recommendations for pilot studies[25], and the results are intended to inform the feasibility of a future validation study rather than support definitive statistical conclusions.

Eligibility criteria included a clinical diagnosis of chronic stroke with upper-limb motor impairment, sufficient cognitive function to understand task instructions, and the ability to provide written informed consent. Initially, 16 individuals were screened for participation. Four were excluded due to severe flaccid paresis ($n = 2$) or cognitive impairment ($n = 2$), resulting in a final sample of 12 post-stroke participants. Since data were collected for both limbs ($n = 24$) and some participants had double hemiparesis, the resulting sample of limbs was $n = 14$ affected and $n = 10$ controls.

All participants provided informed consent before taking part in the study. The research protocol was approved by the institutional ethics committee at the Federal University of Uberlândia, approval number: 39232820.2.0000.5152. All participants provided written informed consent before participation, in accordance with the Declaration of Helsinki. Throughout the assessments, participants were supervised by a licensed physical therapist to ensure safety and compliance.

Feature extraction

To investigate the relationship between motor performance during gameplay and clinical upper-limb function, a set of kinematic features was extracted from each participant's gameplay data. They were processed using custom Python scripts. Although MediaPipe provides 3D coordinates, only the X and Y components were used, as the Z-axis does not provide true depth.

Feature definitions were selected according to the therapeutic focus of the exergame and the quality of camera tracking. Shoulder motion played a minor role in the task, and finger aperture velocity was not extracted because the game did not explicitly require rapid finger movement. The elbow, in contrast, represented the main therapeutic joint and a strong clinical indicator of recovery; thus, both angle and angular velocity were analyzed for this segment. Movements poorly captured from the camera viewpoint were excluded to improve signal reliability. Percentile intervals (e.g., 5th–95th or 25th–75th) were empirically tuned to achieve the best correlation with FMA scores, in line with the regression model's interpretability focus.

The following features were extracted:

Wrist Jitter (Wrist_Jitter). The Wrist Jitter quantifies the instability of movement by measuring frame-to-frame variability in wrist velocity. It is calculated as the standard deviation of the wrist's 2D velocity across the entire recording session.

Maximum Wrist Velocity (Max_Wrist_Vel). Defined as the mean of the top 5% highest instantaneous 2D velocity values of the wrist, calculated between frames.

Average Wrist Velocity (Avg_Wrist_Vel). This feature quantifies the overall magnitude of wrist movement during the session and is defined as the mean of the instantaneous 2D velocities across all frames.

2D Area of Movement (2D_Area). To quantify the spatial extent of wrist movement during each trial, the range of motion was computed from the wrist's (X, Y) coordinates.

To reduce the influence of outliers and tracking noise, the range of motion (ROM) was defined using the 5th and 95th percentiles of the wrist position distribution for each axis, as shown in (1).

$$\mathrm{ROM_x} = P_{95}(x) - P_5(x), \mathrm{ROM_y} = P_{95}(y) - P_5(y) \tag{1}$$

The 2D Area of Movement was then calculated as the product of the horizontal and vertical ROM components, as shown in (2).

$$\mathrm{Area_{2D\ movement}} = \mathrm{ROM_x} \times \mathrm{ROM_y} \tag{2}$$

https://doi.org/10.1038/s41746-026-02383-1

Geometrically, this represents the area of the bounding box that encloses the wrist trajectory. This feature reflects the typical amplitude explored by the participant during the task and serves as an indicator of upper-limb mobility and flexibility.

Total Traveled Distance (Total_Trav_Dist). This feature represents the cumulative path length traveled by the wrist throughout the gameplay session, computed in the two-dimensional plane (X, Y). It was obtained by summing the Euclidean distances between consecutive wrist positions, as shown in (3).

$$D_{\text{total}} = \sum_{i=1}^{n-1} \left| \vec{p_{i+1}} - \vec{p_i} \right|, \text{where} \vec{p_i} = (x_i, y_i) \quad (3)$$

Here, $\vec{p_i}$ is the 2D position of the wrist at frame i, and n is the number of frames. This metric reflects both movement amplitude and frequency, capturing the total excursion of the wrist during the task.

Average Time-to-Target (Avg_Time_Target). For every successful capture, the time-to-target is defined as the interval between the start of the movement segment and the moment the fruit is collected.

Total Score (Score). The total score is the total number of fruits collected at the end of the session. This feature represents overall task performance, as more collected fruits can indicate better performance and coordination.

Average Hand Angle (Avg_Hand_Angle). This feature quantifies the average angle of hand opening and closing throughout the session by computing the flexion angles of each finger. For each frame, the flexion angle of a finger is defined as the angle formed by three anatomical landmarks: the metacarpophalangeal (MCP) joint, the proximal interphalangeal (PIP) joint, and the fingertip (TIP). The angle is calculated as shown in (4),

$$\theta_{\text{finger}} = \cos^{-1}\left(\frac{\vec{a} \cdot \vec{b}}{|\vec{a}||\vec{b}|}\right) \quad (4)$$

where $\vec{a} = \text{PIP} - \text{MCP}$ and $\vec{b} = \text{TIP} - \text{PIP}$.

For each frame, the average flexion angle across the five fingers is computed, as shown in (5).

$$\theta_{\text{mean},i} = \frac{1}{5}\sum_{j=1}^{5} \theta_{j,i} \quad (5)$$

Then, a single global metric is derived by averaging $\theta_{\text{mean,i}}$ across all valid frames. This final value reflects the typical level of finger flexion exhibited during the task and may indicate voluntary control over hand opening and closing.

Average Shoulder Angle (Avg_Shoulder_Angle). The shoulder angle is defined as the angle formed between the trunk (hip-to-shoulder vector) and the upper arm (shoulder-to-elbow vector) at each frame. The average is then computed across all valid frames.

Average Elbow Angle (Avg_Elbow_Angle). This feature is computed frame-by-frame using the angle formed between the upper arm (shoulder-to-elbow vector) and the forearm (elbow-to-wrist vector), extracted from MediaPipe landmarks.

Average Elbow Angular Velocity (Avg_Elbow_Vel). This feature is defined as the mean of the instantaneous angular velocities across all frames.

Maximum Elbow Angular Velocity (Max_Elbow_Vel). This feature is computed as the mean of the top 5% highest instantaneous angular velocities of the elbow observed throughout the session.

This metric serves as a indicator of the participant's ability to generate rapid, forceful elbow motions, often linked to motor control efficiency and residual strength.

Shoulder Angle Range of Motion (ROM_Shoulder_Angle). This feature quantifies the typical angular excursion of the shoulder during the session by computing the interquartile range of the shoulder angle distribution. Specifically, it is defined as the difference between the 75th and 25th percentiles.

Elbow Angle Range of Motion (ROM_Elbow_Angle). This feature is calculated as the interquartile range of the elbow angle distribution, capturing the central 50% of values and excluding extremes.

Hand Angle Range of Motion (ROM_Hand_Angle). This metric estimates the total angular excursion of the hand by capturing the range of average finger flexion angles throughout the session. It is computed as the difference between the 95th and 5th percentiles of the hand flexion angle distribution, as shown in (6):

$$\text{ROM}_{\text{hand}} = P_{95}(\theta_{\text{hand}}) - P_5(\theta_{\text{hand}}) \quad (6)$$

where θ_{hand} denotes the average flexion angle across all fingers for each frame.

Larger values reflect greater capacity for opening and closing the hand, which is relevant to tasks involving grasp and release. Reduced values may indicate spasticity, joint restriction, or lack of voluntary finger extension.

Shoulder-Elbow Correlation (Corr_Shoulder_Elbow). This feature is computed using the Pearson correlation coefficient between the shoulder and elbow angles across all valid frames, as shown in (7):

$$\text{Coord}_{\text{shoulder-elbow}} = \text{corr}(\theta_{\text{shoulder}}, \theta_{\text{elbow}}) \quad (7)$$

This value ranges from -1 (perfectly inverse coordination) to $+1$ (perfectly direct coordination), with values near zero indicating low or inconsistent coupling between the two joints.

Higher correlations may reflect synergistic or stereotyped movement patterns, commonly observed in early or moderate post-stroke recovery phases. Lower values may indicate decoupled joint control, which can emerge with greater motor recovery or impaired coordination. This feature is particularly informative in distinguishing between patients who rely heavily on synergy patterns versus those with more isolated joint control.

Statistical analysis

All analyses were performed using Python 3.11 (SciPy, scikit-learn, and Pandas). To provide an overview of the distribution of each digital feature across severity groups, standardized values (Z-scores) were visualized using a boxplot. This visualization allows for the inspection of descriptive statistics, including the median, standard deviation, interquartile range, and potential outliers.

In addition to statistical feature comparisons, the full sequence of wrist's (X, Y) coordinates was extracted from the gameplay recordings, and overlaid plots were generated by severity group. These plots aimed to qualitatively assess movement range, directionality, and dispersion across groups.

The normality of all variables was assessed using the Shapiro-Wilk test. If the majority of variables demonstrated normal distribution, Pearson's correlation coefficient was used to evaluate the linear association between each feature and the FMA score. Otherwise, Spearman's rank correlation coefficient was used. The significant features were ranked by the correlation coefficient found and presented in a correlation heatmap.

To assess the predictive value of gameplay-derived features, multiple linear regression models were built using FMA scores as the dependent variable and up to five features selected via exhaustive feature selection. All features were standardized using Z-score normalization.

Each limb (affected and non-affected) was treated as an independent observation because they represent distinct functional states with distinct FMA scores and were recorded in separate sessions. Only two participants presented bilateral impairment, and each limb exhibited different degrees of motor deficit. This approach expands the range of FMA scores while maintaining independence between samples. Although not explored here, this design also enables future intra-personal analyses that may control for cognitive or behavioral variability by comparing both limbs of the same

https://doi.org/10.1038/s41746-026-02383-1

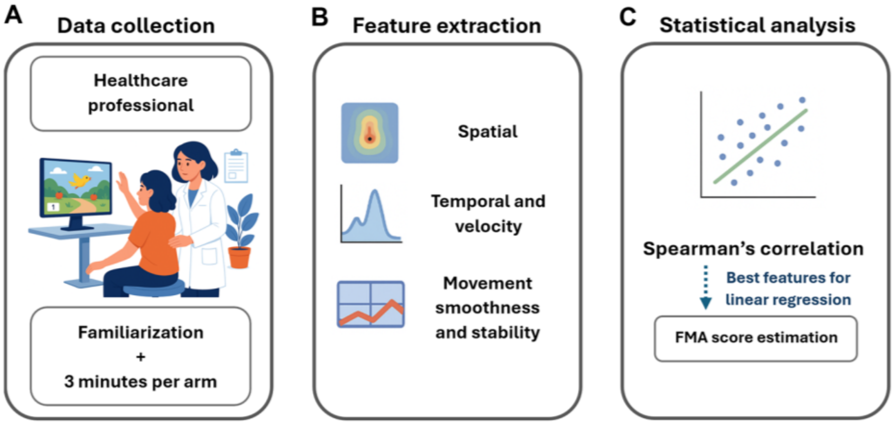


Fig. 8 | Overview of the data processing pipeline. The workflow encompasses three main phases: **A** (left panel) data collection during gameplay with clinical supervision; **B** (central panel) feature extraction of spatiotemporal and kinematic metrics; and **C** (right panel) statistical modeling to examine correlations with upper-limb motor function as measured by the FMA. Source: The authors. Image elements generated with ChatGPT-5; final composition and text edited in Microsoft Word.

individual. Also, this strategy helps reduce variability related to inter-individual cognitive and behavioral differences.

Models were trained and evaluated separately for the affected and total limb groups. Two validation strategies were used: hold-out (50/50 for affected, 70/30 for total) to simulate real-world generalization with approximately 7 participants in the test set, and leave-one-out cross-validation (LOOCV) to maximize data usage and minimize overfitting, especially in small samples.

Model performance was based on the highest Spearman correlation (ρ), lowest root mean square error (RMSE), or highest coefficient of determination (R^2) between predicted and observed FMA values. A scatter plot illustrates the agreement between predicted and true FMA scores for the best model in the affected group.

Given the pilot nature of the dataset and the small number of participants per severity subgroup (particularly mild, $n = 3$), regression models were trained on the total ($n = 24$) and affected-limb ($n = 14$) datasets to ensure broader representation and mitigate class imbalance.

Diagnostic accuracy evaluation

To assess whether predicted FMA values could be used to support clinical classification of motor severity, an exploratory diagnostic accuracy analysis was conducted. The best regression equation identified in the previous analysis was applied to the affected limb subgroup ($n = 14$) data.

The input features were Z-score normalized to ensure consistency with the coefficients used. Real and predicted FMA scores were then categorized into three clinical severity levels based on established thresholds: (i) severe impairment (FMA ≤ 20), (ii) moderate impairment (21 ≤ FMA ≤ 45), (iii) mild impairment (FMA > 45).

A confusion matrix was computed comparing predicted classes to ground truth labels. Subsequently, class-wise precision, recall, F1-score, and accuracy were calculated. Figure 8 shows an overview of the data processing pipeline.

Machine learning analysis

To explore different predictive strategies, four machine learning models were evaluated using a 70/30 hold-out split and Z-score normalization in the total dataset. Performance was assessed with R^2, RMSE, and Spearman's ρ.

The Random Forest was trained on six pre-extracted kinematic features strongly associated with FMA. Using pre-extracted features was justified because Random Forest excels in handling small sample sizes and offers interpretable variable importance, making it well-suited for datasets with limited observations.

In contrast, the 1D-CNN + Bidirectional LSTM model was trained directly on raw time-series hand coordinates (100 Hz). This choice leveraged the ability of deep learning to automatically extract complex temporal patterns from large amounts of sequential data, avoiding the need for manual feature engineering. Predictions from multiple windows were averaged to yield a single FMA estimate per arm.

Finally, Lasso and Elastic Net regressions were applied to the same raw windowed data. These models served as simpler, more interpretable baselines while still allowing direct comparison with the deep learning approach.

Machine learning models were implemented as exploratory baselines to assess feasibility and scalability rather than to provide fully optimized benchmarks. Random Forest was trained on pre-engineered kinematic features due to their stability and transparency in low-sample scenarios, while the other networks were trained on raw time-series data to test the feasibility of end-to-end learning. All models used identical data splits and normalization pipelines to ensure fair comparison. No definitive conclusions regarding model superiority can be drawn due to sample size limitations.

Data availability

The datasets generated and analyzed during the current study are not publicly available at this stage but are available from the corresponding author (J.T.) upon reasonable request. We also intend to make the anonymized dataset publicly accessible through an open research repository following publication. The custom Python scripts used for data preprocessing, feature extraction, and regression analysis are available from the corresponding author upon request.

Received: 17 September 2025; Accepted: 15 January 2026;
Published online: 28 January 2026

References

1. Tsao, C. W. et al. Heart Disease and Stroke Statistics-2022 Update: a report from the american heart association. *Circulation* **145**, e153–e639 (2022).
2. Feigin, V. et al. Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016. *N. Engl. J. Med.* **379**, 2429–2437 (2018).
3. Saposnik, G. & Levin, M. Virtual reality in stroke rehabilitation: a meta-analysis and implications for clinicians. *Stroke* **42**, 1380–1386 (2011).
4. Pyae, A., Luimula, M. & Smed, J. Rehabilitative games for stroke patients. *EAI Endorsed Trans. Serious Games* **1**, e2 (2015).
5. Thomson, K., Pollock, A., Bugge, C. & Brady, M. C. Commercial gaming devices for stroke upper limb rehabilitation: a survey of current practice. *Disabil. Rehabil. Assist Technol.* **11**, 454–461 (2016).
6. Gutiérrez-Pérez, B.-M., Martín-García, A.-V., Murciano-Hueso, A. & de Oliveira Cardoso, A.-P. Use of serious games with older adults: systematic literature review. *Humanit Soc. Sci. Commun.* **10**, 939 (2023).
7. Shahmoradi, L. et al. Virtual reality games for rehabilitation of upper extremities in stroke patients. *J. Bodyw. Mov. Ther.* **26**, 113–122 (2021).
8. de Rooij, I. J. M., van de Port, I. G. L. & Meijer, J.-W. G. Effect of virtual reality training on balance and gait ability in patients with stroke:

systematic review and meta-analysis. *Phys. Ther.* **96**, 1905–1918 (2016).
9. Lohse, K. R., Hilderman, C. G. E., Cheung, K. L., Tatla, S. & Van der Loos, H. F. M. Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. *PLoS ONE* **9**, e93318 (2014).
10. Heller, A. et al. Arm function after stroke: measurement and recovery over the first three months. *J. Neurol. Neurosurg. Psychiatry* **50**, 714–719 (1987).
11. Sunderland, A., Tinson, D., Bradley, L. & Hewer, R. L. Arm function after stroke. An evaluation of grip strength as a measure of recovery and a prognostic indicator. *J. Neurol. Neurosurg. Psychiatry* **52**, 1267–1272 (1989).
12. Nakayama, H., Jørgensen, H. S., Raaschou, H. O. & Olsen, T. S. Recovery of upper extremity function in stroke patients: the copenhagen stroke study. *Arch. Phys. Med. Rehabil.* **75**, 394–398 (1994).
13. Wade, D. T., Langton-Hewer, R., Wood, V. A., Skilbeck, C. E. & Ismail, H. M. The hemiplegic arm after stroke: measurement and recovery. *J. Neurol. Neurosurg. Psychiatry* **46**, 521–524 (1983).
14. Carmona, C. et al. Development and preliminary validity study of a modified version of the upper extremity Fugl-Meyer assessment for use in telerehabilitation. *J. Neurol. Phys. Ther.* **47**, 208–216 (2023).
15. Nie, J. Z., Nie, J. W., Hung, N.-T., Cotton, R. J. & Slutzky, M. W. Portable, open-source solutions for estimating wrist position during reaching in people with stroke. *Sci. Rep.* **11**, 22491 (2021).
16. Song, X., Chen, S., Jia, J. & Shull, P. B. Cellphone-based automated Fugl-Meyer assessment to evaluate upper extremity motor function after stroke. *IEEE Trans. Neural Syst. Rehabil. Eng.* **27**, 2186–2195 (2019).
17. Bai, L. et al. Quantitative assessment of upper limb motion in neurorehabilitation utilizing inertial sensors. *IEEE Trans. Neural Syst. Rehabil. Eng.* **23**, 232–243 (2015).
18. Palani, P., Panigrahi, S., Jammi, S.A. & Thondiyath, A. real-time joint angle estimation using mediapipe framework and inertial sensors. In *2022 IEEE 22nd International Conference on Bioinformatics and Bioengineering (BIBE)* (IEEE, 2022).
19. Rahman, S. et al. AI-driven stroke rehabilitation systems and assessment: a systematic review. *IEEE Trans. Neural Syst. Rehabil. Eng.* **31**, 192–207 (2023).
20. Chen, S. et al. Prediction of the hand function part of the Fugl-Meyer scale after stroke using an automatic quantitative assessment system. *Brain-X* **1**, e26 (2023).
21. Tannus, J., Naves, E. L. M. & Morere, Y. Post-stroke functional assessments based on rehabilitation games and their correlation with clinical scales: a scoping review. *Med Biol. Eng. Comput* **62**, 47–60 (2024).
22. Thomson, K., Pollock, A., Brady, M. & Bugge, C. "The use of commercial gaming devices in upper limb rehabilitation: the experience of stroke survivors. *Int. J. Stroke* **11**, S55 (2016).
23. Google, "MediaPipe | Google for Developers." Accessed: 2025. [Online]. Available: https://developers.google.com/mediapipe.
24. Fugl-Meyer, A. R., Jääskö, L., Leyman, I., Olsson, S. & Steglind, S. A method for evaluation of physical performance. *Scand. J. Rehabil. Med.* **7**, 13–31 (1975).
25. Julious, S. A. Sample size of 12 per group rule of thumb for a pilot study. *Pharm. Stat.* **4**, 287–291 (2005).
26. Song, X., Ding, L., Zhao, J., Jia, J. & Shull, P. Cellphone augmented reality game-based rehabilitation for improving motor function and mental state after stroke. In *2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, (IEEE, 2019).
27. Jiang, Y. et al. A serious game system for upper limb motor function assessment of hemiparetic stroke patients. *IEEE Trans. Neural Syst. Rehabil. Eng.* **31**, 2640–2653 (2023).

Acknowledgements
This work was supported by the National Council for Scientific and Technological Development (CNPq), Brazil (grant numbers 440439/2024-8, 400780/2024-0, and 445256/2024-9), the Coordination for the Improvement of Higher Education Personnel (CAPES), and the Financial Agency for Studies and Projects (FINEP) (grant number 01.24.0116.00). Júlia Tannús is a fellow of CAPES (88887.663238/2022-00). Eduardo Naves is a fellow of CNPq (303614/2023-4).

Author contributions
J.T. conceived and designed the study, performed data analysis, and drafted the manuscript. C.V. contributed to data collection, and patient supervision. E.N. supervised the study, contributed to methodology, and provided critical revisions. All authors reviewed and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to Júlia Tannús.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026

2.3.2 Key Findings and Implications

The reproduced article demonstrates that clinically meaningful motor assessment can be achieved directly during rehabilitation gameplay using simple, interpretable kinematic features extracted from a standard camera setup. The results identify the most promising interpretable features, such as 2D spatial exploration, average finger flexion angles, and inter-joint coordination, which show strong associations with motor function. These features can be used individually as estimators or combined within a multiple linear regression model to estimate clinical scores, as demonstrated by the prediction of the FMA.

The integration of assessment into gameplay eliminates the need for separate evaluation procedures, enabling continuous, real-time monitoring of motor function while reducing clinical workload.

Within the context of this thesis, this article consolidates the proposed framework by demonstrating that a fully sensor-free, low-cost, and interpretable system can simultaneously support rehabilitation and automated clinical assessment. It directly addresses the limitations identified in earlier phases and establishes the foundation for scalable telerehabilitation applications.

2.4 Scientific contributions

This doctoral research advances the field of post-stroke rehabilitation by proposing and validating a low-cost, sensor-free, and interpretable framework for simultaneous motor rehabilitation and assessment based on exergames.

First, this work provides a systematic consolidation of the state of the art regarding the use of rehabilitation games and virtual environments for motor assessment. Through a structured literature review, it demonstrates that gameplay-derived metrics can achieve clinically meaningful correlations with standardized scales, supporting their use as digital biomarkers for motor function evaluation.

Second, this thesis introduces a low-cost vision-based motion tracking system capable of estimating upper-limb kinematics with clinically acceptable accuracy using only a standard RGB camera. The proposed geometric approach demonstrates that reliable motion capture can be achieved without relying on expensive sensors, contributing to the accessibility and scalability of rehabilitation technologies.

Finally, the research advances by proposing a fully integrated framework in which rehabilitation and assessment occur simultaneously during gameplay. This eliminates the need for separate clinical evaluation sessions and enables continuous, real-time monitoring of motor performance.

A key contribution of this work is the demonstration that simple, interpretable kinematic features are sufficient to estimate clinical motor scores with high accuracy. The results show that features related to hand configuration, spatial exploration, and joint coordination are strongly associated with the FMA and can be combined into lightweight predictive models.

Additionally, this work introduces a design paradigm shift toward hardware-independent rehabilitation systems. The adoption of AI-based tracking enables the complete removal of external physical dependencies, significantly improving usability, scalability, and real-world applicability.

Finally, this thesis contributes to the concept of embedded digital biomarkers, in which clinically meaningful metrics are extracted passively during therapy. This approach supports high-frequency, non-intrusive monitoring of patient progress and lays the foundation for scalable telerehabilitation and remote assessment systems.

2.5 Methodological limitations

Despite the promising results, this study presents limitations and further validation difficulties that should be acknowledged. First, the use of the FMA as the reference clinical scale required considerable therapist time and necessitated patient travel to the testing location, hindering large-scale testing, as all data collection had to be performed uniformly by the same clinician.

The primary limitation is the relatively small sample size, consisting of 12 post-stroke participants and 24 evaluated upper limbs. Although this is consistent with pilot and feasibility studies, it restricts the statistical power of the analyses and limits the generalizability of the results.

Additionally, the study adopts a cross-sectional design, evaluating participants in a single session. As a result, it does not capture longitudinal changes in motor recovery, limiting the ability to assess responsiveness to rehabilitation over time or to evaluate the system’s sensitivity to clinical progression.

Another limitation concerns the controlled experimental conditions under which data were collected. Although efforts were made to simulate realistic usage, factors such as camera positioning, lighting conditions, and participant posture were partially standardized. This may not fully reflect the variability encountered in home-based or unsupervised environments.

Also, the use of 2D vision-based tracking introduces inherent constraints, particularly regarding depth estimation and sensitivity to occlusions. While the AI-based approach enables hardware independence, it may exhibit reduced accuracy compared to depth cameras or multi-sensor systems in certain scenarios.

Another unresolved limitation concerns generalizability: the regression equation derived in this study is tailored to the current exergame and cannot be directly transferred to other game designs or movement contexts. However, the variables identified are interpretable and could be further studied to be used as independent indicators of functional monitoring.

3 CONCLUSION

This thesis presented the development and validation of a low-cost, vision-based framework for post-stroke rehabilitation and automated motor assessment, integrating therapeutic interaction and clinical evaluation into a single exergame system.

The research was structured in three progressive phases. First, a literature review established that gameplay-derived metrics can correlate with clinical scales, supporting their use as digital biomarkers. Second, a vision-based tracking system demonstrated that accurate kinematic data can be obtained using accessible technologies. Finally, an integrated AI-driven framework showed that these data can be used to estimate clinical motor scores directly during gameplay.

The results demonstrate that clinically relevant motor assessment can be performed using simple, interpretable features extracted from standard camera input, without the need for specialized hardware. The proposed system achieved high predictive performance in estimating FMA scores and successfully classified patients according to motor impairment severity.

A key finding of this work is that model simplicity and interpretability can be more advantageous than complexity in rehabilitation contexts. Lightweight regression models not only achieved strong performance but also provided transparency and clinical relevance, which are essential for real-world adoption.

Another important contribution is the transition toward hardware-independent rehabilitation systems, enabling broader accessibility and scalability. By embedding assessment into gameplay, the proposed framework reduces clinical workload, increases patient engagement, and enables continuous monitoring of recovery.

Overall, this thesis demonstrates that it is feasible to transform rehabilitation games into clinically meaningful assessment tools, contributing to the advancement of digital health solutions for stroke rehabilitation.

3.1 Future work

Future research should focus on addressing the limitations identified in this study and further advancing the proposed framework.

A primary direction is the validation of the system with larger and more diverse populations, including different stages of stroke recovery and broader demographic variability. This will improve the robustness and generalizability of the predictive models.

Longitudinal studies are also necessary to evaluate the system's ability to track motor recovery over time and to assess its sensitivity to clinical changes. This is essential for establishing the framework as a reliable tool for continuous monitoring and outcome evaluation.

Another important direction is the incorporation of multi-rater clinical assessments, enabling evaluation of inter-rater reliability and strengthening the clinical validity of the results.

Further investigation into feature selection and personalization may also enhance predictive performance, allowing models to adapt to individual patient characteristics and rehabilitation profiles.

Finally, the integration into telerehabilitation platforms, including remote data transmission, clinician dashboards, and real-time feedback systems, represents a step toward clinical translation and large-scale adoption.

3.2 Articles published during the doctorate studies

This subsection presents the contributions to scientific literature published by the author during the doctoral studies (2022-2026).

3.2.1 Articles related to the thesis

- TANNUS, Julia; VALENTINI, Caroline; NAVES, Eduardo. AI-driven low-cost rehabilitation exergame as a lightweight framework for stroke assessment. *npj digital medicine*, 2026. https://doi.org/10.1038/s41746-026-02383-1
- TANNUS, Julia *et al.* Low-cost vision-based 3D elbow tracking for post-stroke rehabilitation: development and pilot evaluation of a serious game. *IEEE transactions on neural systems and rehabilitation engineering*, v. 34, p. 1–1, 2025. http://doi.org/10.1109/TNSRE.2025.3591104
- TANNUS, Julia; NAVES, Eduardo Lázaro Martins; MORERE, Yann. Post-stroke functional assessments based on rehabilitation games and their correlation with clinical scales: a scoping review. *Medical & biological engineering & computing*, v. 62, p. 47, 2023. http://doi.org/10.1007/s11517-023-02933-9
- TANNUS, Julia *et al.* Correlation between the Fugl-Meyer Assessment and data from a virtual reality game for remote evaluation of upper-limb function in post-stroke rehabilitation: a case study. In: *HANDICAP 2024 – 13th IFRATH conference on assistive technologies*, 2024, Paris, p. 115–120.

3.2.2 Articles published in partnership with the research group

- ROCHA, Danilo Santos *et al.* Effectiveness of telerehabilitation in the treatment of shoulder injuries: a systematic review of randomized controlled trials. *Archives of rehabilitation research and clinical translation*, v. 1, p. 100553, 2025. http://doi.org/10.1016/j.arrct.2025.100553
- DE SOUZA MIGUEL, Guilherme Fernandes *et al.* Proposal of a game streaming-based framework for a telerehabilitation system. *Multimedia tools and applications*, v. 83, p. 33333, 2023. http://doi.org/10.1007/s11042-023-16741-8
- MARQUES, Isabela Alves *et al.* Virtual reality and serious game therapy for post-stroke individuals: a preliminary study with a humanized rehabilitation approach protocol. *Complementary therapies in clinical practice*, v. 49, p. 101681, 2022. http://dx.doi.org/10.33448/rsd-v10i6.15489
- ALVES, Camille Marques *et al.* A new human–machine interface for the rehabilitation of individuals with Parkinson's disease. In: *HANDICAP 2022 – 12th IFRATH conference on assistive technologies*, 2022, Paris, 2022. p. 81–86.

BIBLIOGRAPHY

ADAMS, Richard J *et al.* Upper Extremity Function Assessment Using a Glove Orthosis and Virtual Reality System. *OTJR: occupation, participation and health*, v. 39, n. 2, p. 81–89, abr. 2019. https://doi.org/10.1177/1539449219829862

AHMAD, Mohd Azzuan *et al.* Virtual Reality Games as an Adjunct in Improving Upper Limb Function and General Health among Stroke Survivors. *International journal of environmental research and public health*, v. 16, n. 24, dez. 2019. https://doi.org/10.3390/ijerph16245144

ALLEGUE, Dorra Rakia *et al.* Rehabilitation of Upper Extremity by Telerehabilitation Combined With Exergames in Survivors of Chronic Stroke: Preliminary Findings From a Feasibility Clinical Trial. *JMIR rehabilitation and assistive technologies*, v. 9, n. 2, p. e33745, 2022. https://doi.org/10.2196/33745

ANTHES, C *et al.* State of the art of virtual reality technology. *2016 IEEE aerospace conference*, 2016. p. 1–19. https://doi.org/10.1109/AERO.2016.7500674

BAI, Lu *et al.* Quantitative Assessment of Upper Limb Motion in Neurorehabilitation Utilizing Inertial Sensors. *IEEE transactions on neural systems and rehabilitation engineering*, v. 23, n. 2, p. 232–243, 2015. https://doi.org/10.1109/TNSRE.2014.2369740

BARTEIT, Sandra *et al.* Augmented, mixed, and virtual reality-based head-mounted devices for medical education: systematic review. *JMIR serious games*, v. 9, n. 3, p. e29080, 2021. https://doi.org/10.2196/29080

BERNHARDT, Julie *et al.* Moving Rehabilitation Research Forward: Developing Consensus Statements for Rehabilitation and Recovery Research. *Neurorehabilitation and Neural Repair*, v. 31, n. 8, p. 694–698, 2017. https://doi.org/10.1177/1545968317724290

BOWMAN, D A; MCMAHAN, R P. Virtual Reality: How Much Immersion Is Enough? *Computer*, v. 40, n. 7, p. 36–43, 2007. https://doi.org/10.1109/MC.2007.257

BURDEA, Grigore *et al.* Robotic Table and Serious Games for Integrative Rehabilitation in the Early Poststroke Phase: Two Case Reports. *JMIR rehabilitation and assistive technologies*, v. 9, n. 2, p. e26990, abr. 2022. https://doi.org/10.2196/26990

CAO, Z *et al.* OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE transactions on pattern analysis and machine intelligence*, v. 43, n. 1, p. 172–186, 2021. https://doi.org/10.1109/TPAMI.2019.2929257

CARMONA, Carolina *et al.* Development and Preliminary Validity Study of a Modified Version of the Upper Extremity Fugl-Meyer Assessment for Use in Telerehabilitation. *Journal of neurologic physical therapy: JNPT*, v. 47, n. 4, p. 208–216, out. 2023. https://doi.org/10.1097/NPT.0000000000000447

CHEN, Shugeng *et al.* Prediction of the hand function part of the Fugl-Meyer scale after stroke using an automatic quantitative assessment system. *Brain-X*, v. 1, n. 3, p. e26, 2023. https://doi.org/10.1002/brx2.26

CIPRESSO, Pietro *et al.* The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature. *Frontiers in psychology*, v. 9, 2018. https://doi.org/10.3389/fpsyg.2018.02086

DUNCAN, P W; PROPST, M; NELSON, S G. Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. *Physical therapy*, v. 63, n. 10, p. 1606–1610, out. 1983. https://doi.org/10.1093/ptj/63.10.1606

FABBRIZIO, Antonio *et al.* Smart devices for health and wellness applied to Tele-exercise: An overview of new trends and technologies such as IoT and AI. *Healthcare (Basel, Switzerland)*, v. 11, n. 12, 2023. https://doi.org/10.3390/healthcare11121805

FEIGIN, Valery L *et al.* Global, regional, and national burden of stroke and its risk factors, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. *The Lancet Neurology*, v. 23, n. 10, p. 973–1003, 1 out. 2024. https://doi.org/10.1016/S1474-4422(24)00369-7

FUGL-MEYER, Axel R *et al.* A method for evaluation of physical performance. *Scandinavian journal of rehabilitation medicine*, v. 7, n. 1, p. 13–31, 1975. https://doi.org/10.2340/1650197771331

GLADSTONE, David J; DANELLS, Cynthia J; BLACK, Sandra E. The Fugl-Meyer Assessment of Motor Recovery after Stroke: A Critical Review of Its Measurement Properties. *Neurorehabilitation and neural repair*, v. 16, n. 3, p. 232–240, 1 set. 2002. https://doi.org/10.1177/154596802401105171

GOOGLE. *MediaPipe | Google for Developers*. Available at: <https://developers.google.com/mediapipe>. Acessed on: jan. 18, 2026.

GU, Yuexing *et al.* A Review of Hand Function Rehabilitation Systems Based on Hand Motion Recognition Devices and Artificial Intelligence. *Brain sciences*, v. 12, n. 8, 2022. https://doi.org/10.3390/brainsci12081079

GUO, Yan *et al.* A Survey of the State of the Art in Monocular 3D Human Pose Estimation: Methods, Benchmarks, and Challenges. *Sensors*, v. 25, n. 8, 2025. https://doi.org/10.3390/s25082409

JIANG, Yuanbo *et al.* A Serious Game System for Upper Limb Motor Function Assessment of Hemiparetic Stroke Patients. *IEEE transactions on neural systems and rehabilitation engineering*, v. 31, p. 2640–2653, 2023. https://doi.org/10.1109/TNSRE.2023.3281408

KIM, Won-Seok *et al.* Upper Extremity Functional Evaluation by Fugl-Meyer Assessment Scoring Using Depth-Sensing Camera in Hemiplegic Stroke Patients. *PloS one*, v. 11, n. 7, p. e0158640, 2016. https://doi.org/10.1371/journal.pone.0158640

KIPER, Pawel *et al.* Reinforced feedback in virtual environment for rehabilitation of upper extremity dysfunction after stroke: preliminary data from a randomized controlled trial. *BioMed research international*, v. 2014, p. 752128, 2014. https://doi.org/10.1155/2014/752128

KLEIN, Luan *et al.* Assessing the Reliability of AI-Based Angle Detection for Shoulder and Elbow Rehabilitation. *Optimization, learning algorithms and applications*, 2024. p. 3–18. https://doi.org/10.1007/978-3-031-53036-4_1

LANGHORNE, Peter; BERNHARDT, Julie; KWAKKEL, Gert. Stroke rehabilitation. *The Lancet*, v. 377, n. 9778, p. 1693–1702, 2011. https://doi.org/10.1016/S0140-6736(11)60325-5

LEE, Seunghee; LEE, Yang-Soo; KIM, Jonghyun. Automated Evaluation of Upper-Limb Motor Function Impairment Using Fugl-Meyer Assessment. *IEEE transactions on neural systems and rehabilitation engineering*, v. 26, n. 1, p. 125–134, 2018. https://doi.org/10.1109/TNSRE.2017.2755667

LOHSE, Keith R *et al.* Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. *PloS one*, v. 9, n. 3, p. e93318, 2014. https://doi.org/10.1371/journal.pone.0093318

LUGARESI, Camillo *et al.* MediaPipe: A Framework for Building Perception Pipelines. *arXiv:1906.08172*, 2019. https://doi.org/10.48550/arXiv.1906.08172

MAGGIO, Maria Grazia *et al.* The overlooked role of exergames in cognitive-motor neurorehabilitation: a systematic review. *npj Digital Medicine*, v. 8, n. 1, p. 419, 2025. https://doi.org/10.1038/s41746-025-01843-4

MANSER, Patrick *et al.* Beyond “just” fun: The role of exergames in advancing health promotion and disease prevention. *Neuroscience & biobehavioral reviews*, v. 176, p. 106260, 2025. https://doi.org/10.1016/j.neubiorev.2025.106260

MARTIN, Seth S *et al.* 2024 Heart Disease and Stroke Statistics: A Report of US and Global Data From the American Heart Association. *Circulation*, v. 149, n. 8, p. e347–e913, 20 fev. 2024. https://doi.org/10.1161/CIR.0000000000001247

MASMOUDI, Mostefa *et al.* Assessing the effectiveness of virtual reality serious games in post-stroke rehabilitation: a novel evaluation method. *Multimedia tools and applications*, v. 83, n. 12, p. 36175–36202, 2024. https://doi.org/10.1007/s11042-023-17980-5

MATHIOWETZ, Virgil *et al.* Adult Norms for the Box and Block Test of Manual Dexterity. *The American journal of occupational therapy*, v. 39, p. 386–391, 1 jul. 1985. https://doi.org/10.5014/ajot.39.6.386

MÉNDEZ, Ángela *et al.* Influence of Demographic and Clinical Factors on Perceived Usability, Presence, Flow, Competence, Pleasant and Unpleasant Sensations, and Utility During Interaction with Virtual Reality Games for Motor and Cognitive Rehabilitation: An Observational Study in Patients with Stroke and Traumatic Brain Injury. *Games for health journal*, 2025. https://doi.org/10.1177/2161783X251370421

MICHAEL, David; CHEN, Sande. Serious games: games that educate, train, and inform. *Boston, USA: Thomson Course Technology*, 2006.

NEWELL, Alejandro; YANG, Kaiyu; DENG, Jia. Stacked Hourglass Networks for Human Pose Estimation. *Computer vision – ECCV 2016*, 2016. p. 483–499. https://doi.org/10.1007/978-3-319-46484-8_29

NIE, Jeffrey Z *et al.* Portable, open-source solutions for estimating wrist position during reaching in people with stroke. *Scientific reports*, v. 11, n. 1, p. 22491, 2021. https://doi.org/10.1038/s41598-021-01805-2

NOROUZI-GHEIDARI, Nahid *et al.* Feasibility, Safety and Efficacy of a Virtual Reality Exergame System to Supplement Upper Extremity Rehabilitation Post-Stroke: A Pilot Randomized Clinical Trial and Proof of Principle. *International journal of environmental research and public health*, v. 17, n. 1, 2020. https://doi.org/10.3390/ijerph17010113

OLIVEIRA, Sérgio *et al.* Recovering through play: Studying the effects of collaborative Virtual Reality serious games for stroke rehabilitation through a human-centered design methodology. *Computers & graphics*, v. 134, p. 104501, 2026. https://doi.org/10.1016/j.cag.2025.104501

PAGE, Stephen J; LEVINE, Peter; HADE, Erinn. Psychometric Properties and Administration of the Wrist/Hand Subscales of the Fugl-Meyer Assessment in Minimally Impaired Upper Extremity Hemiparesis in Stroke. *Archives of physical medicine and rehabilitation*, v. 93, n. 12, p. 2373- 2376.e5, 2012. https://doi.org/10.1016/j.apmr.2012.06.017

PALANI, Poongavanam *et al.* Real-time Joint Angle Estimation using Mediapipe Framework and Inertial Sensors. *2022 IEEE 22nd international conference on bioinformatics and bioengineering*, 2022. p. 128–133. https://doi.org/10.1109/BIBE55377.2022.00035

PARK, Yu-Hyung; LEE, Chi-Ho; LEE, Byoung-Hee. Clinical usefulness of the virtual reality-based postural control training on the gait ability in patients with stroke. *Journal of exercise rehabilitation*, v. 9, n. 5, p. 489–494, 2013. https://doi.org/10.12965/jer.130066

POURMAND, Ali *et al.* Emerging Utility of Virtual Reality as a Multidisciplinary Tool in Clinical Medicine. *Games for health journal*, v. 6, n. 5, p. 263–270, 2017. https://doi.org/10.1089/g4h.2017.0046

POURMAND, Ali *et al.* Virtual reality as a clinical tool for pain management. *Current pain and headache reports*, v. 22, p. 1–6, 2018. https://doi.org/10.1007/s11916-018-0708-2

RAHMAN, Sejuti *et al.* AI-Driven Stroke Rehabilitation Systems and Assessment: A Systematic Review. *IEEE transactions on neural systems and rehabilitation engineering*, v. 31, p. 192–207, 2023. https://doi.org/10.1109/TNSRE.2022.3219085

RODRIGUEZ-DE-PABLO, C *et al.* Validating ArmAssist Assessment as outcome measure in upper-limb post-stroke telerehabilitation. *Annual international conference of the IEEE engineering in medicine and biology society*, 2015. p. 4623–4626. https://doi.org/10.1109/EMBC.2015.7319424

RODRIGUEZ-DE-PABLO, Cristina *et al.* Development of computer games for assessment and training in post-stroke arm telerehabilitation. *Annual international conference of the IEEE engineering in medicine and biology society*, v. 2012, p. 4571–4574, 2012. https://doi.org/10.1109/EMBC.2012.6346984

SÁNCHEZ-HERRERA-BAEZA, Patricia *et al.* The Impact of a Novel Immersive Virtual Reality Technology Associated with Serious Games in Parkinson's Disease Patients on Upper Limb Rehabilitation: A Mixed Methods Intervention Study. *Sensors*, v. 20, n. 8, 2020. https://doi.org/10.3390/s20082168

SLATER, Mel; SANCHEZ-VIVES, Maria V. Enhancing Our Lives with Immersive Virtual Reality. *Frontiers in Robotics and AI*, v. 3, 2016. https://doi.org/10.3389/frobt.2016.00074

SONG, X; DING, L; *et al.* Cellphone Augmented Reality Game-based Rehabilitation for Improving Motor Function and Mental State after Stroke. *2019 IEEE 16th international conference on wearable and implantable body sensor networks*, 2019. p. 1–4. https://doi.org/10.1109/BSN.2019.8771093

SONG, X; CHEN, S; *et al.* Cellphone-Based Automated Fugl-Meyer Assessment to Evaluate Upper Extremity Motor Function After Stroke. *IEEE transactions on neural systems and rehabilitation engineering*, v. 27, n. 10, p. 2186–2195, 2019. https://doi.org/10.1109/TNSRE.2019.2939587

TANNUS, Julia *et al.* Low-Cost Vision-Based 3-D Elbow Tracking for Post-Stroke Rehabilitation: Development and Pilot Evaluation of a Serious Game. *IEEE transactions on neural systems and rehabilitation engineering*, v. 33, 2025. https://doi.org/10.1109/TNSRE.2025.3591104

TANNUS, Julia; VALENTINI, Caroline; NAVES, Eduardo. AI-driven low-cost rehabilitation exergame as a lightweight framework for stroke assessment. *npj digital medicine*, 2026. https://doi.org/10.1038/s41746-026-02383-1

TANNUS, Julia; NAVES, Eduardo; DE SÁ, Angela. Computer vision devices for tracking gross upper limb movements in post-stroke rehabilitation. *Research, society and development*, v. 10, n. 6, p. e57910616143, 2021. https://doi.org/10.33448/rsd-v10i6.16143

TANNUS, Julia; NAVES, Eduardo L M; MORERE, Yann. Post-stroke functional assessments based on rehabilitation games and their correlation with clinical scales: A scoping review. *Medical & biological engineering & computing*, v. 62, n. 1, p. 47–60, 2024. https://doi.org/10.1007/s11517-023-02933-9

THOMSON, Katie *et al.* Commercial gaming devices for stroke upper limb rehabilitation: a survey of current practice. *Disability and rehabilitation: assistive technology*, v. 11, n. 6, p. 454–461, 2016. https://doi.org/10.3109/17483107.2015.1005031

THOMSON, Katie *et al.* Commercial gaming devices for stroke upper limb rehabilitation: The stroke survivor experience. *Journal of rehabilitation and assistive technologies engineering*, v. 7, 2020. https://doi.org/10.1177/2055668320915381

VEERBEEK, Janne Marieke *et al.* What Is the Evidence for Physical Therapy Poststroke? A Systematic Review and Meta-Analysis. *PloS one*, v. 9, n. 2, p. e87987-, 2014. https://doi.org/10.1371/journal.pone.0087987

WEBER, Lynne M *et al.* Immersive virtual reality mirror therapy for upper limb recovery following stroke: A pilot study. *American journal of physical medicine & rehabilitation*, v. 98, n. 9, p. 783, 2019. https://doi.org/10.1097/PHM.0000000000001190

WINSTEIN, Carolee J *et al.* Guidelines for Adult Stroke Rehabilitation and Recovery. *Stroke*, v. 47, n. 6, p. e98–e169, 2016. https://doi.org/10.1161/STR.0000000000000098

WOLF, Steven L *et al.* The EXCITE Trial: Attributes of the Wolf Motor Function Test in Patients with Subacute Stroke. *Neurorehabilitation and neural repair*, v. 19, n. 3, p. 194–205,2005. https://doi.org/10.1177/1545968305276663

YEH, Shih-Ching *et al.* An innovative virtual reality system for mild cognitive impairment: diagnosis and evaluation. *2012 IEEE-EMBS conference on biomedical engineering and sciences*, 2012, p. 23–27. https://doi.org/10.1109/IECBES.2012.6498023

ANNEX I – FUGL-MEYER ASSESSMENT FOR UPPER LIMBS

FMA-UE PROTOCOL	Rehabilitation Medicine, University of Gothenburg

FUGL-MEYER ASSESSMENT UPPER EXTREMITY (FMA-UE) Assessment of sensorimotor function

ID:
Date:
Examiner:

Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S: The post-stroke hemiplegic patient. A method for evaluation of physical performance. Scand J Rehabil Med 1975, 7:13-31.

A. UPPER EXTREMITY, sitting position

I. Reflex activity	**none**	**can be elicited**
Flexors: biceps and finger flexors (at least one)	0	2
Extensors: triceps	0	2
Subtotal I (max 4)		

II. Volitional movement within synergies, without gravitational help			**none**	**partial**	**full**
Flexor synergy: Hand from contralateral knee to ipsilateral ear. From extensor synergy (shoulder adduction/ internal rotation, elbow extension, forearm pronation) to flexor synergy (shoulder abduction/ external rotation, elbow flexion, forearm supination).	Shoulder	retraction	0	1	2
		elevation	0	1	2
		abduction (90°)	0	1	2
		external rotation	0	1	2
	Elbow	flexion	0	1	2
	Forearm	supination	0	1	2
Extensor synergy: Hand from ipsilateral ear to the contralateral knee	Shoulder	adduction/internal rotation	0	1	2
	Elbow	extension	0	1	2
	Forearm	pronation	0	1	2
		Subtotal II (max 18)			

III. Volitional movement mixing synergies, without compensation		**none**	**partial**	**full**
Hand to lumbar spine hand on lap	cannot perform or hand in front of ant-sup iliac spine	0		
	hand behind ant-sup iliac spine (without compensation)		1	
	hand to lumbar spine (without compensation)			2
Shoulder flexion 0°- 90° elbow at 0° pronation-supination 0°	immediate abduction or elbow flexion	0		
	abduction or elbow flexion during movement		1	
	flexion 90°, no shoulder abduction or elbow flexion			2
Pronation-supination elbow at 90° shoulder at 0°	no pronation/supination, starting position impossible	0		
	limited pronation/supination, maintains starting position		1	
	full pronation/supination, maintains starting position			2
	Subtotal III (max 6)			

IV. Volitional movement with little or no synergy		**none**	**partial**	**full**
Shoulder abduction 0 - 90° elbow at 0° forearm neutral	immediate supination or elbow flexion	0		
	supination or elbow flexion during movement		1	
	abduction 90°, maintains extension and pronation			2
Shoulder flexion 90° - 180° elbow at 0° pronation-supination 0°	immediate abduction or elbow flexion	0		
	abduction or elbow flexion during movement		1	
	flexion 180°, no shoulder abduction or elbow flexion			2
Pronation/supination elbow at 0° shoulder at 30°- 90° flexion	no pronation/supination, starting position impossible	0		
	limited pronation/supination, maintains start position		1	
	full pronation/supination, maintains starting position			2
	Subtotal IV (max 6)			

V. Normal reflex activity assessed only if full score of 6 points is achieved in part IV; compare with the unaffected side		**hyper**	**lively**	**normal**
Biceps, triceps, finger flexors	2 of 3 reflexes markedly hyperactive	0		
	1 reflex markedly hyperactive or at least 2 reflexes lively		1	
	maximum of 1 reflex lively, none hyperactive			2
	Subtotal V (max 2)			
	Total A (max 36)			

GÖTEBORGS UNIVERSITET

B. WRIST support may be provided at the elbow to take or hold the starting position, no support at wrist, check the passive range of motion prior testing		**none**	**partial**	**full**
Stability at 15° dorsiflexion elbow at 90°, forearm pronated shoulder at 0°	less than 15° active dorsiflexion dorsiflexion 15°, no resistance tolerated maintains dorsiflexion against resistance	0	1	2
Repeated dorsifexion / volar flexion elbow at 90°, forearm pronated shoulder at 0°, slight finger flexion	cannot perform volitionally limited active range of motion full active range of motion, smoothly	0	1	2
Stability at 15° dorsiflexion elbow at 0°, forearm pronated slight shoulder flexion/abduction	less than 15° active dorsiflexion dorsiflexion 15°, no resistance tolerated maintains dorsiflexion against resistance	0	1	2
Repeated dorsifexion / volar flexion elbow at 0°, forearm pronated slight shoulder flexion/abduction	cannot perform volitionally limited active range of motion full active range of motion, smoothly	0	1	2
Circumduction elbow at 90°, forearm pronated shoulder at 0°	cannot perform volitionally jerky movement or incomplete complete and smooth circumduction	0	1	2
	Total B (max 10)			

C. HAND support may be provided at the elbow to keep 90° flexion, no support at the wrist, compare with unaffected hand, the objects are interposed, active grasp		**none**	**partial**	**full**
Mass flexion from full active or passive extension		0	1	2
Mass extension from full active or passive flexion		0	1	2
GRASP				
a. Hook grasp flexion in PIP and DIP (digits II-V), extension in MCP II-V	cannot be performed can hold position but weak maintains position against resistance	0	1	2
b. Thumb adduction 1-st CMC, MCP, IP at 0°, scrap of paper between thumb and 2-nd MCP joint	cannot be performed can hold paper but not against tug can hold paper against a tug	0	1	2
c. Pincer grasp, opposition pulpa of the thumb against the pulpa of 2-nd finger, pencil, tug upward	cannot be performed can hold pencil but not against tug can hold pencil against a tug	0	1	2
d. Cylinder grasp cylinder shaped object (small can) tug upward, opposition of thumb and fingers	cannot be performed can hold cylinder but not against tug can hold cylinder against a tug	0	1	2
e. Spherical grasp fingers in abduction/flexion, thumb opposed, tennis ball, tug away	cannot be performed can hold ball but not against tug can hold ball against a tug	0	1	2
	Total C (max 14)			

D. COORDINATION/SPEED, sitting, after one trial with both arms, eyes closed, tip of the index finger from knee to nose, 5 times as fast as possible		**marked**	**slight**	**none**
Tremor	at least 1 completed movement	0	1	2
Dysmetria	pronounced or unsystematic slight and systematic no dysmetria	0	1	2
		≥ 6s	**2 - 5s**	**< 2s**
Time start and end with the hand on the knee	6 or more seconds slower than unaffected side 2-5 seconds slower than unaffected side less than 2 seconds difference	0	1	2
	Total D (max 6)			

TOTAL A-D (max 66)	

GÖTEBORGS UNIVERSITET