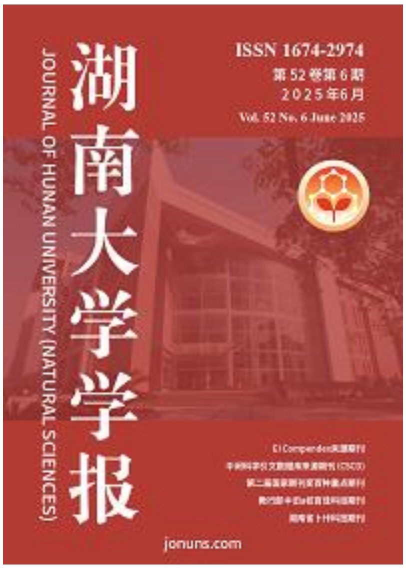

Journal of Hunan University

(Natural Sciences)

Vol. 52 No. 6

June 2025

Available online at

https://jonuns.com

Open Access Article https://doi.org/10.55463/issn.1674-2974.52.6.3

Virtual Intelligent Agent for Hand Therapy Support

Robinson Jiménez-Moreno[1], Anny Astrid Espitia-Cubillos[1*], Mateo Pulido-Aponte[2]

[1]Associated Professors of Engineering Faculty, Universidad Militar Nueva Granada, Bogotá, Colombia

[2]Research assistant, Engineering Faculty, Universidad Militar Nueva Granada, Bogotá, Colombia

* Corresponding author: anny.espitia@unimilitar.edu.co

Article History:

Received: May 29, 2025
Revised: June 23, 2025
Accepted: July 2, 2025
Published: July 30, 2025

Abstract: This paper presents the design and implementation of an artificial-intelligence-driven virtual therapist that guides repetitive hand-opening and-closing exercises for rehabilitation—accessible to people of any age, from any location, on any webcam-equipped computer. The agent integrated varies techniques for its quasi-natural interaction with the user, to determinate the correct advance during the therapy is supported on recurrent networks of long-short term memory (LSTM) is used. The network allows the processing of features obtained from hand movement samples, which contain information on positions, distances and angles, as well as time-controlled image captures for specific periods of time. These features are obtained using MediaPipe, an open-source framework developed by Google that allows discrimination of characteristic points of the hand. The proposed development focuses on performing a classification in time series samples of the opening and closing movement of the hand in real time, to define whether the user is performing correctly or not a basic exercise for strengthening hand joints. As contribution to agent's finds, a large language model (LLM) is integrated for the analysis of the training results (hand exercise session) and the generation of a detailed report. The system is integrated into a web application developed using two frameworks that demarcate the creation of the agent. The design and user experience were developed using React, the management and data processing was performed using FastAPI. It is possible to evidence the easy use of the agent as a therapeutic assistant and an accurate progress report of the rehabilitation therapy, where according to an accuracy of 88% in learning of the LSTM, it accurately discriminates the movement in hand physical therapy.

Keywords: artificial intelligence, assistive robotics, physical therapy, long and short term memory (LSTM) networks, React agent, recurrent neural network.

支持手部治疗的虚拟智能代理

摘要： 本文介绍了一种基于人工智能的代理的开发，目的是实现一种技术工具，以支持物理手部治疗，该治疗包括重复的开合动作，可以在任何地方由任何年龄段的人士使用，只要他们有带摄像头的计算机即可。该代理集成了多种技术，以实现其与用户的类自然互动，治疗过程中的正确进展是基于长短期记忆（LSTM）递归网络的支持。。该网络可以处理手部运动样本中获得的特征，这些样本包含位置、距离和角度信息，以及一定时间段内的时间控制图像捕获。这些特征是通过 MediaPipe 获得的，MediaPipe 是谷歌开发的一个开源框架，可以识别手部的特征点。拟议的开发重点是对手部开合运动的实时时间序列样本进行分类，以确定用户是否正确地进行了加强手部关节的基本练习。该系统集成了一个大型语言模型（LLM），用于分析训练结果（手工练习）和生成详细报告。该系统集成在一个网络应用程序中，使用两个框架开发，这两个框架划分了代理的创建。设计和用户体验使用 React 开发，管理和数据处理使用 FastAPI。可以证明的是，该代理可作为治疗助手轻松使用，并可提供准确的康复治疗进度报告，其中 LSTM 在学习中准确判别手部物理治疗中的运动的准确率为 88%。

关键词： 人工智能、辅助机器人、物理治疗、长短期记忆（LSTM）网络、React 代理、递归神经网络。

1. Introduction

Physical hand therapy is required when there are injuries resulting from accidents, illness and/or the passage of time, where traditional rehabilitation methods often take a long time and rely on hospital resources [1]. Conventional therapy is done in a hospital or at home with the accompaniment of professional therapists, which is not always affordable for developing countries [2], due to limitations such as transporting the patient to the physical therapist's office or vice versa, economic difficulties, problems for travel by distance or vehicular mobility, lack of companion in the case of infants or elderly people, lack of availability of agenda, among others. Whereas virtual therapists based on intelligent agents offer a technological tool safe and easy-to-use, designed to be used anywhere and at any time, this development would provide a solution to the need for a particular therapy from remote ways.

Healthcare systems have evolved through the acquisition of bio-signals by advanced sensors [3] and through the development of so-called smart homes [4]. Where intelligent systems with the ability to act autonomously and proactively in their environment to achieve specific goals are leading to a technological revolution with future impact [5]. A fundamental part of this are agent-based systems, used in applications such as anomaly detection in smart grids [6], voice assistants in smart homes [7], intelligent management of water resources [8], monitoring of intelligent vehicles [9] and various applications in smart industry based on the Internet of Things (IoT) [10], thus developing the integral concept of smart cities, which are centered on multi-agent systems [11].

The integration of IoT devices based on artificial intelligence is generating solutions in what are called intelligent healthcare systems [12] [13]. This gives rise to a new concept, the virtual therapist that today contributes to scenarios such as psychiatry [14], speech disorders [15], therapies in cancer patients [16] and assistive and rehabilitative technologies in patients with autism [17]. The adoption of virtual therapeutic techniques [18], for example, art-based [19] or prosthetic [20] can be developed from home. Where the use of technological tools such as virtual environments or even the use of robotic exoskeletons [21] are part of what is now called intelligent therapy [22].

With respect to hand therapy, the analysis of their movements using artificial intelligence is feasible [23]. In [24] they point out artificial intelligence as an area of interest in hand therapy for children with cerebral palsy for personalization, adaptation and use of data. There are developments such as a controller with AI technology for the adaptation of a library of therapeutic games personalized according to arm and hand impairments that offers outcome tracking [25]. Also, in [26] authors present a web platform, based on computer vision, containing digital serious games designed as hand therapy exercises. In [27] a virtual hand rehabilitation system that integrates a web application with a deep learning model is presented, which in turn complements feature extraction from a ResNeXt50 network with sequential analysis of LSTM data to recognize gestures, compares the patient's video with a reference one, and features an exercise performance tracker, a chatbot, a scheduler for appointments with physical therapists, and a SQLite database that records each exercise.

In line with this technological evolution in the field of therapy-oriented health care, this paper presents the design and implementation of a virtual intelligent agent for hand therapies. So that by training an LSTM network [28], whose learning is based on the movement of a hand opening and closing therapy exercise, it is possible to estimate the current position and pose of the user in front of the virtual agent by means of a recognition algorithm called MediaPipe [29], already used in movement rehabilitation applications [30] together with LSTM networks [31].

The proposed agent guides the user in the development of the exercise by means of voice prompts, and by making a record of the task execution and generating a concept by means of an LLM about the performance in the hand therapy exercise. This type of tool facilitates access to physical hand therapy from anywhere, giving greater independence to the patient, who can determine at what time of the day he/she carries out the activity with timely feedback favoring hand rehabilitation through a therapy directed in real time. The main contribution within the development intelligent agent that monitors the rehabilitation process through the integration of artificial intelligence techniques that also acts as an expert in the area by the LLM implemented.

2. Methodology

Designing a virtual intelligent agent oriented to physical therapy of the hand involving a focus on applied investigation with experimental validation. Given that the quasi-natural interaction between the user and the agent is searched and must be established the criteria of interaction through adecuate prompts for the LLM y for the recognition of the therapy action based on the hand movement.

To develop an intelligent agent to guide physical therapy focused on hand opening and closing, and to validate and evaluate the progress in the exercise by the user in front of a webcam, an artificial intelligence algorithm is used, based on a short-long term memory network (LSTM). This network has a RNN (recurrent neural network) architecture that uses sequential information such as that provided by the hand movement in the opening and closing sequence, which works well with images and videos, as is the interaction that is exposed with the proposed agent and the user through the image captured by webcam.

The development of the project includes five stages. The first stage corresponds to the initial design of the agent with its components. In the second stage, the selection of tools to be integrated into the agent is carried out. Thirdly, the necessary requirements for each component of the agent are defined. In the fourth stage, which is the most important, the detailed design of each component is carried out. Finally, in the fifth stage, the agent is tested for validation.

3. Results

3.1. Initial agent design

In the initial design of the agent, three fundamental parts are established as components: the frontend for visualization and interaction with the user; the backend for processing and report generation; and the LSTM-based model for exercise classification, capturing movement by camera, as shown in Figure 1.

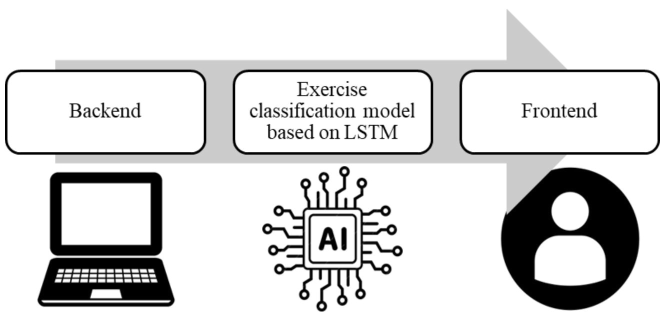

Figure 1. Initial agent design (Source: developed by the authors)

3.2. Tool definition

For the definition of the tools to be integrated in the system, we considered the animation and voice reproduction of the agent, describing the performance of the physical therapy by text, detecting and following the user's hands live and generating coherent text in the final report of each therapy. Table 1 lists each tool used, with the task it will perform in the agent, its description and the component to which it corresponds.

Table 1. Tools used for agent design (Source: developed by the authors)

Tool	Task	Description	Component
React Native	Animation and voice playback	Open-source JavaScript library that allows to build interactive and efficient user interfaces	FrontEnd
Fast API	Text description of performance	High-performance web framework for building APIs with Python	BackEnd
Media Pipe	Hand detection and tracking	Google's cross-platform open-source framework for building custom machine learning modeling solutions for live and streaming multimedia content [23]	Model for exercise classification
Open AI API	Consistent text generation	Interface that allows developers to access advanced artificial intelligence models and integrate them into their applications	FrontEnd

3.3. Delimitation of requirements

The basic requirements necessary for the integration of the agent to perform the classification of a specific exercise for the strengthening of joints in hand therapies were delimited, which are presented in Figure 2 according to the three basic components of the design: backend, Frontend and AI model by LSTM.

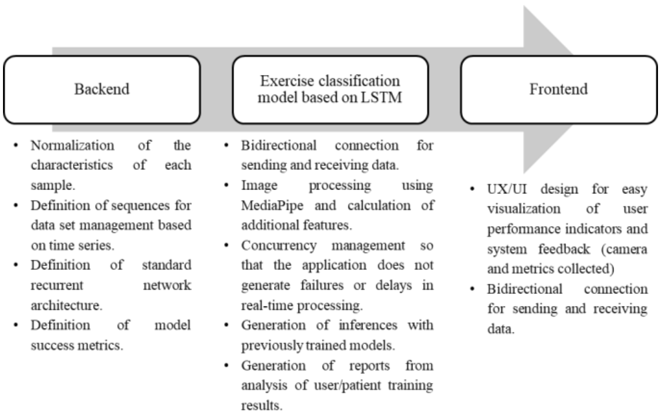

Figure 2. Requirements by component (Source: developed by the authors)

3.4. Detailed design of each component

3.4.1. Frontend

A simple but functional design was defined to allow the user to understand both the indications for use and performance of the therapeutic exercise, as well as the processed signals and feedback from the system, the sketch is presented in Figure 3. This template determines what the user will see on the screen.

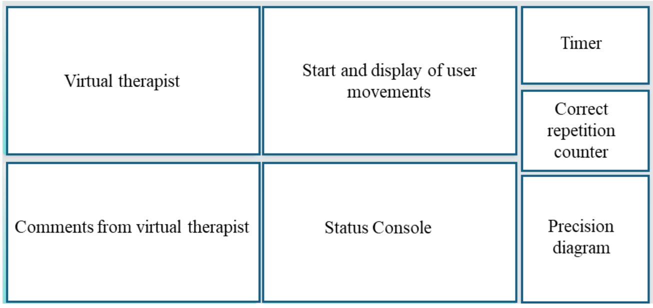

Figure 3. FrontEnd block diagram (Source: developed by the authors)

Figure 4 and according to the template illustrates the final user interaction interface (frontend). The upper left side shows the virtual therapist (named Terry) presenting basic animations of mouth and eye movements, Terry gives feedback to the patient verbally using the predictions generated by each hand movement sample. At the lower left is a text box that replicates the verbal instructions given to the user. In the upper central part is "Entrenar" button that starts the therapy session and captures the user's camera. If the system is correctly capturing the hand movement samples, indicators of the position of the joints should appear. In the lower central part of the screen is the Log Console, which shows the level of prediction, inference time and CPU and memory usage in the computer where the agent is running. In the upper right part is the time counter, an indicator that allows the user to know the duration of the training in minutes and seconds; below it is the repetition counter that shows the number of repetitions correctly performed by the user; and finally in the lower right part is pie chart that illustrates the precision in exercise execution and is updated in real time during the session.

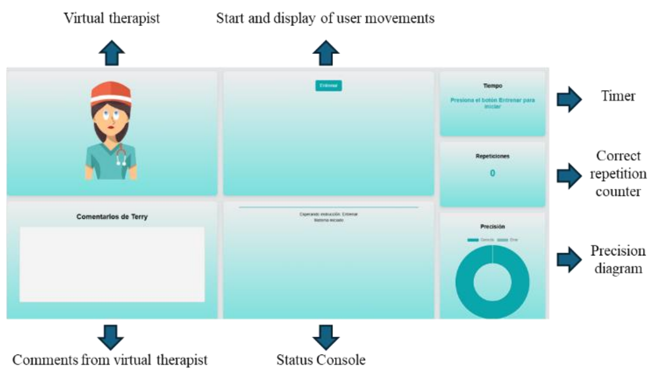

Figure 4. FrontEnd (Source: developed by the authors)

3.4.2. Backend

The server-side application contains several functionalities:

1. Real-time, two-way communication via websocket from frontend to backend and vice versa, allowing management of the data displayed and processed.

2. Transmission of images captured with webcam and

training states (exercise session) by the frontend. These states allow management of when to store samples for exercise inference, as represented in Figure 5.

3. Reading images and validating states for image processing, extracting features using MediaPipe and calculating additional features, such as distances and angles of hand movement.

4. Depending on the value of the training states, the samples are stored in CSV files according to the flowchart in Figure 6.

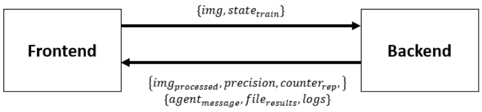

Figure 5. Frontend - Backend Communication (Source: developed by the authors)

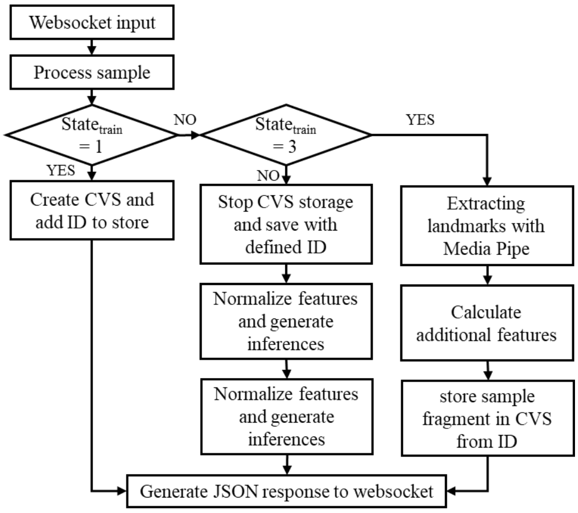

Figure 6. Training states flowchart (Source: developed by the authors)

3.4.3. LSTM-based model

From CSV files containing information on captured samples, when the user performs an exercise from start (closing the hand) to finish (opening the hand), inferences are generated using an LSTM-based model that works with the sequence represented in Figure 7.

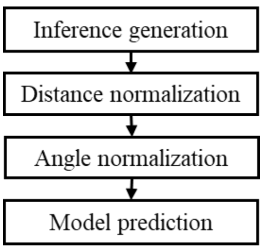

Figure 7. LSTM network processing sequence (Source: developed by the authors)

The recognition of the opening and closing action is based on the training of the Bi-LSTM network, for which a basic architecture is defined for its compilation, considering the use case of binary classification and time series prediction in sequential data, such as those obtained from hand movements. For this purpose, the Bi-LSTM network architecture illustrated in Figure 8 was defined.

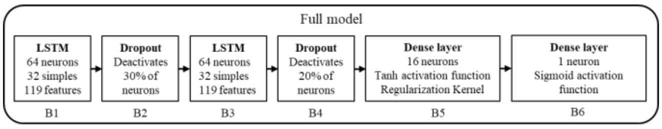

Figure 8. LSTM Architecture (Source: developed by the authors)

Below, each layer of the network architecture used by the blocks (B) outlined in Figure 8 is explained:

B1. LSTM 64: This is a long-short-term memory layer with 64 neural units that allows for capturing long-term dependencies of the temporal sequence.

B2. Dropout 30%: This randomly deactivates 30% of the neurons during training to avoid overfitting, improving generalization. The 30% was selected because results with higher values were tested, but this decreased the model's accuracy.

B3. LSTM 32: This is a long-short-term memory layer with 32 neural units that allows for capturing short-term dependencies learned from the temporal sequence.

B4. Dropout 20%: This randomly deactivates 20% of the neurons during training, reducing overfitting. Since there are fewer neurons in the previous layer, this value is decreased to maintain the previously adjusted weights.

B5. 16-neuron dense layer: Uses the hyperbolic tangent activation function, useful for modeling nonlinear relationships in the data, since the data are not uniform across samples.

B6. 1-neuron dense layer: This is the output layer and contains a single neuron because problem has been modeled binary, generating a value between 0 and 1 that represents probability of belonging to following classes:

- Correct exercise (> 0.75).
- Incorrect exercise (< 0.75).

This value of 0.75 is because during training, there was a larger data set for samples labeled with correct exercises than for incorrect samples. Therefore, when calculating the proportion of weights in the set for each type of label, a value of 1.5 was obtained for the label "0" (incorrect) and 0.75 for the label "1" (correct). This balances the model's predictions.

To optimize the training phase, the type of normalization to be used was defined. For distances referenced by the characteristic points of the hand using MediaPipe, the minimum-maximum method was used (equation 1) since the measurements are relative and are given in pixels.

$$X_N = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

As indicated, feature extraction is performed from the landmarks found with MediaPipe, as can be seen in Figure 9.

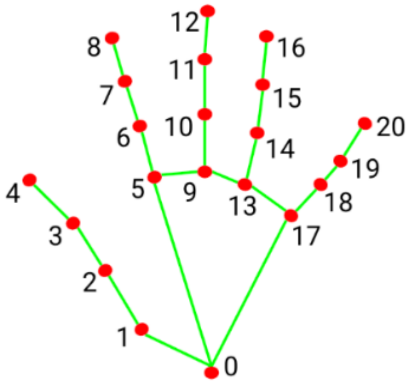

Figure 9. Landmarks hand (Source: [23])

These landmarks are automatically positioned upon detecting a hand, allowing positions to be discriminated and measurements to be calculated so that the model can optimally find patterns in its training and prediction phase. To achieve this, distances are calculated for each of the hand's fingers, as well as the angles between each landmark, generating more data for model training.

While increasing the amount of elements in the training set, the variability of the compartment tends to reduce that the trends of training and validation data sets tend to perform a very similar behavior, so it can be deduced that by increasing the size of the data set probably the model can generate better predictions, however for the case of use of binary classification the behavior obtained is sufficient.

Finally, a PDF file is generated summarizing the execution time, total expected repetitions, total correctly performed repetitions, training accuracy, details for each repetition, and a general analysis. For the report analysis, the Gemini-1.5-flash large language model (LLM) analyzes the results of the completed training. A physiotherapy expert role analyzes the patient's positive and areas of improvement, and generates a detailed analysis with recommendations using the prompt:

input = You are an expert hand therapist. Analyze the following information from a hand therapy training session:
- Number of exercises:{counter_attempt}
- Number of exercises performed correctly: {attempt_success}
- Table of predictions by attempts:
- Predictions made: {list_predictions}
- Attempts made: {list_attempts}
Generate a very concise analysis of the training session, highlighting the positive and areas of improvement, and their impact on the patient.
Express yourself as a healthcare professional.

3.5. Proofs

Based on the above, the network was trained by varying the number of samples (*N*) in the data set, taking a constant of 20% of the data in each case for validation and maintaining a maximum limit of 50 epochs with 8 subgroups (batch size). Table 2 illustrates the comparison of the training sessions varying from 88 to 678 data, where accuracy is obtained from 82% to 88%. In the case of the best network, it corresponds to 424 data with an accuracy of 88.24%, precision of 91.67%.

Table 2. LSTM training comparison (Source: developed by the authors)

N	Accuracy	Precision	AUC	Recall
88	0.8235	0.8000	0.7667	1.0000
170	0.8824	0.9048	0.9267	0.9048
255	0.8627	0.9189	0.9089	0.8947
424	0.8824	0.9167	0.9400	0.9167
678	0.8235	0.8588	0.9082	0.8588

The training of the model, despite using a CPU, that is, without parallel processing on the GPU, took less than 1 minute even using the largest number of samples (678). This is because the calculations were performed with the measurements extracted from positions, distances and angles, that is, images were not processed but their descriptors.

It was observed that by increasing the number of samples in the training set for the LSTM network, accuracy, precision, recall and area under the curve - ROC (AUC) presented positive indicators with trends to 1, however, this can mean that the model is on adjustment, so it is convenient to add stages to the DRPOUT model which allows in most cases to avoid said overhap. posterior metrics. When analyzing the behavioral graphs of the loss and accuracy function of each training (see table 2), it can be seen that for the first training the model is on adjusting because the training and validation trends are opposite and in the case of the loss function tends to increase the set of validation and otherwise, in the accuracy phase the trend of the validation set has to decrease, in addition to showing a high variability.

A validation environment was established for the use of the agent with four progressions in hand therapy, simulating gradual progress, where very low, low, medium and high therapy progress was defined. Figure 10 illustrates the intelligent agent use environment for the test scenarios.

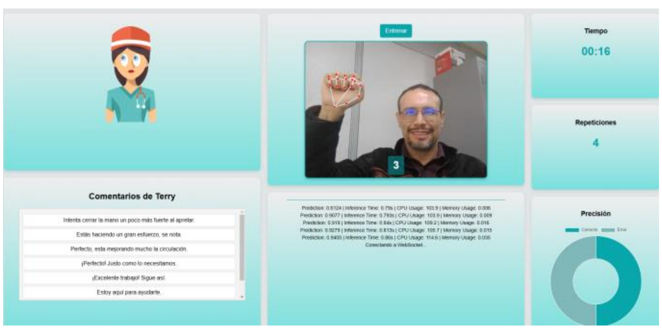

Figure 10. Progression test environment in therapy (Source: developed by the authors)

Table 3 relates the upper part of the reports generated by the intelligent agent for each case of progression where the precision levels in the year range from 14.28% to 100%, indicating in which repetition the exercise failed.

Table 3. PDF reports (Source: developed by the authors)

<table>
<tr><th>Therapy progress</th><th>Report</th></tr>
<tr><td>Very low</td><td>Tiempo de ejecución: 60s.
Repeticiones esperadas: 7.
Repeticiones realizadas correctamente: 1.
Precisión del entrenamiento: 14.286.
Detalle por cada repetición:
<table><tr><td>Repetición</td><td>Análisis</td></tr><tr><td>1</td><td>Incorrecto</td></tr><tr><td>2</td><td>Correcto</td></tr><tr><td>3</td><td>Incorrecto</td></tr><tr><td>4</td><td>Incorrecto</td></tr><tr><td>5</td><td>Incorrecto</td></tr><tr><td>6</td><td>Incorrecto</td></tr><tr><td>7</td><td>Incorrecto</td></tr></table></td></tr>
<tr><td>Low</td><td>Tiempo de ejecución: 60s.
Repeticiones esperadas: 5.
Repeticiones realizadas correctamente: 1.
Precisión del entrenamiento: 20.0.
Detalle por cada repetición:
<table><tr><td>Repetición</td><td>Análisis</td></tr><tr><td>1</td><td>Incorrecto</td></tr><tr><td>2</td><td>Correcto</td></tr><tr><td>3</td><td>Incorrecto</td></tr><tr><td>4</td><td>Incorrecto</td></tr><tr><td>5</td><td>Incorrecto</td></tr></table></td></tr>
<tr><td>Half</td><td>Tiempo de ejecución: 60s.
Repeticiones esperadas: 7.
Repeticiones realizadas correctamente: 4.
Precisión del entrenamiento: 57.143.
Detalle por cada repetición:
<table><tr><td>Repetición</td><td>Análisis</td></tr><tr><td>1</td><td>Correcto</td></tr><tr><td>2</td><td>Incorrecto</td></tr><tr><td>3</td><td>Correcto</td></tr><tr><td>4</td><td>Incorrecto</td></tr><tr><td>5</td><td>Correcto</td></tr><tr><td>6</td><td>Incorrecto</td></tr><tr><td>7</td><td>Correcto</td></tr></table></td></tr>
<tr><td>High</td><td>Tiempo de ejecución: 60s.
Repeticiones esperadas: 7.
Repeticiones realizadas correctamente: 7.
Precisión del entrenamiento: 100.0.
Detalle por cada repetición:
<table><tr><td>Repetición</td><td>Análisis</td></tr><tr><td>1</td><td>Correcto</td></tr><tr><td>2</td><td>Correcto</td></tr><tr><td>3</td><td>Correcto</td></tr><tr><td>4</td><td>Correcto</td></tr><tr><td>5</td><td>Correcto</td></tr><tr><td>6</td><td>Correcto</td></tr><tr><td>7</td><td>Correcto</td></tr></table></td></tr>
</table>

Finally, through the large language model (LLM) based on Gemini with an expert role in physiotherapy, the following types of analysis are generated, as examples of established validation cases.

The patient has performed a total of seven hand therapy exercises, of which he has carried out four correctly. This implies a 57% success percentage, which is a good sign, since it indicates that the patient is managed to progress in their rehabilitation.

More extensive analysis of the model even generates additional recommendations against possible moods of the patient according to the following case:

Hand therapy training showed discouraging results. Of five exercises, only one was done correctly. Although there is success in the second attempt, the incorrect execution in the remaining four exercises indicates an urgent need for adjustment in the therapeutic strategy. The patient presents poor understanding or execution of the exercises. This can negatively impact the functional recovery of the hand. An exhaustive revaluation of the program is required, including a detailed review of the instruction, the modification of the difficulty of the exercises, and the consideration of possible limiting factors of the patient's ability, such as pain or lack of strength. More basic and progressive exercises must be prioritized to ensure a solid base before advancing to more complex exercises. The motivation and understanding of the patient should also be evaluated and reinforced.

4. Discussion

The designed agent stands out from previous developments that virtually support the development of hand therapy [6], [29], [32] where intelligent therapy lacks a local autonomous guide in the patient's location, for this case the agent (Terry), in turn incorporates the use of a LLM that provides with the generation of automatic reports and real -time feedback both by voice and by text to the user.

The benefits inherent in the development of intelligent agents operating locally for therapy and supported by LLM's, either under executable code or web access, facilitate and reduce the costs of access to therapy for patients, clearly allow for reduced analysis times of therapy progressions and optimize the use of hospital rehabilitation infrastructures.

5. Conclusion

During the agent's development, aspects were evidenced that feedback decision making in the design of this type of intelligent agent allows it to achieve conclusions based on the management of concurrences for proper real-time data management generated delays in development because the signals were not being processed with latency times less than 1 second, so it was decided angular measures.

It is evidenced that the orientation and progress within the simulation of rehabilitation and respective verbal feedback from the intelligent agent, aiding the user in the development of a guided and quantified exercise, which allows them to act upon so that they can adjust their progress and validate it in real time through the agent answers', obtaining historical reports useful for the medical evaluation by experts. Denoting the application potential in therapy of these types of development and how it is implemented on different schemes of therapy that do not require physical interaction between user and therapist.

Declarations

Author Contributions

Conceptualization, formal analysis and writing—review and editing Jiménez-Moreno R. and Espitia-Cubillos A.; methodology, validation, investigation supervision, project administration, funding acquisition, and data curation Jiménez-Moreno R.; Software, writing—original draft preparation and visualization

Pulido Aponte M. All authors have read and agreed to the published version of the manuscript.

Funding

Product derived from the research project titled "Diseño de un modelo de interacción humano robot mediante algoritmos de aprendizaje profundo" INV-ING-3971 financed by the vice-rector for research of the Universidad Militar Nueva Granada, year 2024.

Acknowledgements

The authors thank the Universidad Militar Nueva Granada for the time and resources available for the development of this article where they are associate professors.

Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committe of the Universidad Militar Nueva Granada (project INV-ING-3971, date of approval: 18/01/2024).

Conflicts of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

References

[1] D. N. SOUMIS, and N. D. TSELIKAS. Serious Game for Hand Therapy Assessment, in *13th International Conference on Modern Circuits and Systems Technologies (MOCAST),* Sofia, Bulgaria, 2024. https://doi.org/10.1109/MOCAST61810.2024.10615316

[2] R. K. MEGALINGAM, S. K. MANOHARAN, S. M. MOHANDAS, C. P. K. REDDY, E. N. P. N. V. K. VIJAY, and Y D. CHANDRIKA. Wearable Hand Orthotic Device for Rehabilitation: Hand Therapy with Multi-Mode Control and Real-Time Feedback. *Applied Sciences*, 2023, 13(6): 3976. https://doi.org/10.3390/app13063976

[3] G. DIRACO. Chapter 1 - Advanced sensors for smart healthcare: an introduction, in *Advanced Sensors for Smart Healthcare*, T. A. Nguyen, Ed., Elsevier, 2025, 1-27. https://doi.org/10.1016/B978-0-443-24790-3.00002-8

[4] Z. LUO, J. PENG, X. ZHANG, H. JIANG, R. YIN, Y. TAN, and M. LV. Optimal scheduling of smart home energy systems: A user-friendly and adaptive home intelligent agent with self-learning capability. *Advances in Applied Energy*, 2024, 15: 100182. https://doi.org/10.1016/j.adapen.2024.100182

[5] S. HOSSEINI, and H. SEILANI. The role of agentic AI in shaping a smart future: A systematic review. *Array*, 2025, 26: 100399. https://doi.org/10.1016/j.array.2025.100399

[6] J. WANG. Multi agent system based smart grid anomaly detection using blockchain machine learning model in mobile edge computing network. *Computers and Electrical Engineering*, 2025, 121: 109825. https://doi.org/10.1016/j.compeleceng.2024.109825

[7] M. HUBERT, S. KAZEMI, M. HUBERT, A. CARUGATI, and M. M. MARIANI. The exploration of users' perceived value from personalization and virtual conversational agents to enable a smart home assemblage– A mixed method approach. *International Journal of Information Management*, 2025, 80: 102850. https://doi.org/10.1016/j.ijinfomgt.2024.102850

[8] H. MANCY, N. E. GHANNAM, A. ABOZEID, and A. I. TALOBA. Decentralized multi-agent federated and reinforcement learning for smart water management and disaster response. *Alexandria Engineering Journal*, 2025, 126: 8-29. https://doi.org/10.1016/j.aej.2025.04.033

[9] M.-H. YEN, N. MISHRA, W.J. LUO, and C.E. LIN. A Novel Proactive AI-Based Agents Framework for an IoE-Based Smart Things Monitoring System with Applications for Smart Vehicles. *Computers, Materials and Continua*, 2025, 82(2): 1839-1855. https://doi.org/10.32604/cmc.2025.060903

[10] Y. WANG, S. SU, and Y. WANG. Attention-augmented multi-agent collaboration for Smart Industrial Internet of Things task offloading. *Internet of Things*, 2025, 31: 101572. https://doi.org/10.1016/j.iot.2025.101572

[11] L. V. D. NASCIMENTO, and J. P. M. D. OLIVEIRA. A multi-agent architecture for context sources integration in smart cities. *Future Generation Computer Systems*, 2025, 172: 107862. https://doi.org/10.1016/j.future.2025.107862

[12] G. SINGH. Wearable IoT (w-IoT) artificial intelligence (AI) solution for sustainable smart-healthcare. *International Journal of Information Management Data Insights*, 2025, 5(1): 100291. https://doi.org/10.1016/j.jjimei.2024.100291

[13] B. RAVIKRISHNA, M. E. SENO, M. RAPARTHI, R. R. YELLU, S. ALSUBAI, A. K. DUTTA, A. AZIZ, D. ABDURAKHIMOVA, and J. BHOLA. Artificial intelligence probabilities scheme for disease prevention data set construction in intelligent smart healthcare scenario. *SLAS Technology*, 2024, 29(4): 100164. https://doi.org/10.1016/j.slast.2024.100164

[14] S. M. SCHUELLER, C. STILES-SHIELDS, and L. YAROSH. Online Treatment and Virtual Therapists in Child and Adolescent Psychiatry. *Child and Adolescent Psychiatric Clinics of North America*, 2017, 26(1): 1-12. https://doi.org/10.1016/j.chc.2016.07.011

[15] Y. P. P. CHEN, C. JOHNSON, P. LALBAKHSH, T. CAELLI, G. DENG, D. TAY, S. ERICKSON, P. BROADBRIDGE, A. E. REFAIE, W. DOUBE, and M. E. MORRIS. Systematic review of virtual speech therapists for speech disorders. *Computer Speech & Language*, 2016, 37: 98-128. https://doi.org/10.1016/j.csl.2015.08.005

[16] A. MARKS, A. GARBATINI, V. P. KIMBERLY HIEFTJE, V. WESER, and C. S. F. FERNANDES. Use of Immersive Virtual Reality Spaces to Engage Adolescent and Young Adult Patients With Cancer in Therapist-Guided Support Groups: Protocol for a pre-post study. *JMIR Research Protocols*, 2023, 12: 48761. https://doi.org/10.2196/48761

[17] J. MILLS, and O. DUFFY. Speech and Language Therapists' Perspectives of Virtual Reality as a Clinical Tool for Autism: Cross-Sectional Survey. *JMIR Rehabilitation and Assistive technologies*, 2025, 12: 63235. https://doi.org/10.2196/63235

[18] A. MILOFF, P. CARLBRING, W. HAMILTON, G. ANDERSSON, L. REUTERSKIÖLD, and P. LINDNER. Measuring Alliance Toward Embodied Virtual Therapists in

the Era of Automated Treatments With the Virtual Therapist Alliance Scale (VTAS): Development and Psychometric Evaluation. *Journal of Medical Internet Research*, 2020, 22(3): 16660. https://doi.org/10.2196/16660
[19] K. HAN, and S. KONG. Exploring factors affecting art therapists' intention to embrace virtual reality using UTAUT1. *The Arts in Psychotherapy*, 2025, 92: 102260. https://doi.org/10.1016/j.aip.2025.102260
[20] S. G. ROZEVINK, B. MAAS, A. MURGIA, R. M. BONGERS, and C. K. V. D. SLUIS. Therapists' and prosthesis users' assessments of a virtual reality environment designed for upper limb prosthesis control training. *Journal of Hand Therapy*, 2025. https://doi.org/10.1016/j.jht.2025.01.004
[21] T. AHMED, E. L.-G. M. R. ISLAM, I. WANG, and M. RAHMAN. Robot-aided Rehabilitation with SREx: A Smart Robotic Exoskeleton for Reducing Therapists' Physical Stress. *Archives of Physical Medicine and Rehabilitation*, 2022, 103(12): 184-185. https://doi.org/10.1016/j.apmr.2022.08.936
[22] T. H. NGUYEN, M. ELIOS, A. M. FITZPATRICK, S. M. NYENHUIS, and W. PHIPATANAKUL. Optimizing Pediatric Asthma Management: Implementation and Evaluation of AIR and SMART Therapies. *The Journal of Allergy and Clinical Immunology: In Practice*, 2025. https://doi.org/10.1016/j.jaip.2025.05.004
[23] W. BAE, N. JUNG, Y. JUNG, and K. KAM. Analysis of the applicability of an artificial intelligence technology to classify hand movements of people with or without limited hand movement, using AI-motion SW: An observational study. *Edelweiss Applied Science and Technology*, 2025, 9(3): 2930–2940. https://doi.org/10.55214/25768484.v9i3.5889
[24] T. V. P. CISNEROS, A. BRONS, B. KRÖSE, B. SCHOUTEN, and G. LUDDEN. Playfulness and New Technologies in Hand Therapy for Children With Cerebral Palsy: Scoping Review. *JMIR Serious Games*, 2023, 11(1): 44904. https://doi.org/10.2196/44904
[25] G. BURDEA, N. KIM, K. POLISTICO, A. KADARU, N. GRAMPUROHIT, D. ROLL, and F. DAMIANI. Assistive game controller for artificial intelligence-enhanced telerehabilitation post-stroke. *Assistive Technology*, 2019, 33(3): 117–128. https://doi.org/10.1080/10400435.2019.1593260
[26] D. N. SOUMIS, and N. D. TSELIKAS. A Web-Based Platform for Hand Rehabilitation Assessment. *Big Data and Cognitive Computing*, 2025, 9(3): 52. https://doi.org/10.3390/bdcc9030052
[27] P. MANIAR, G. VEDPATHAK, and S. FRANCIS. VTherapy: Deep Learning-based Hand Rehabilitation System. In *2nd International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI)*, Wardha, India, 2024. https://doi.org/10.1109/IDICAIEI61867.2024.10842730
[28] J. GONZALEZ, and W. YU. Non-linear system modeling using LSTM neural networks. *IFAC-PapersOnLine*, 2018, 51(13): 485-489. https://doi.org/10.1016/j.ifacol.2018.07.326
[29] W. SIMOES, L. REIS, C. ARAUJO, and J. M. JR. Accuracy Assessment of 2D Pose Estimation with MediaPipe for Physiotherapy Exercises. *Procedia Computer Science*, 2024, 251: 446-453. https://doi.org/10.1016/j.procs.2024.11.132
[30] A. LATRECHE, R. KELAIAIA, A. CHEMORI, and A. KERBOUA. Reliability and validity analysis of MediaPipe-based measurement system for some human rehabilitation motions. *Measurement*, 2023, 214: 112826. https://doi.org/10.1016/j.measurement.2023.112826
[31] S. RANI, S. MASOOD, and D. R. RIZVI. A Bi-LSTM based approach for Compensation detection during robotic stroke rehabilitation therapy. *Procedia Computer Science*, 2025, 258: 3145-3154. https://doi.org/10.1016/j.procs.2025.04.572
[32] A. WOLFF, G. WEINSTOCK-ZLOTNICK, and J. GORDON. SSc management – In person appointments and remote therapy (SMART): A framework for management of chronic hand conditions. *Journal of Hand Therapy*, 2014, 27(2): 143-151. https://doi.org/10.1016/j.jht.2013.11.004

参考文:

[1] D. N. SOUMIS, and N. D. TSELIKAS. 严肃游戏用于手部治疗评估，第 13 *届现代电路与系统技术国际会议 (MOCAST)*，保加利亚索非亚，2024. https://doi.org/10.1109/MOCAST61810.2024.10615316
[2] R. K. MEGALINGAM, S. K. MANOHARAN, S. M. MOHANDAS, C. P. K. REDDY, E. N. P. N. V. K. VIJAY, and Y D. CHANDRIKA. 用于康复的可穿戴手部矫形装置：具有多模式控制和实时反馈的手部治疗。*应用科学*, 2023, 13(6): 3976. https://doi.org/10.3390/app13063976
[3] G. DIRACO. 第 1 章 - 智能医疗的先进传感器：简介 *智能医疗的先进传感器*，T.A. Nguyen，编辑，爱思唯尔, 2025, 1-27. https://doi.org/10.1016/B978-0-443-24790-3.00002-8
[4] Z. LUO, J. PENG, X. ZHANG, H. JIANG, R. YIN, Y. TAN, and M. LV. 智能家居能源系统的优化调度：具有自学习能力的用户友好且自适应的家居智能代理。*应用能源的进展*，2024, 15: 100182. https://doi.org/10.1016/j.adapen.2024.100182
[5] S. HOSSEINI, and H. SEILANI. 代理人工智能在塑造智能未来中的作用：系统评价。*大批*, 2025, 26: 100399. https://doi.org/10.1016/j.array.2025.100399
[6] J. WANG. 基于多代理系统的智能电网异常检测，使用移动边缘计算网络中的区块链机器学习模型。*计算机与电气工程*，2025, 121: 109825. https://doi.org/10.1016/j.compeleceng.2024.109825
[7] M. HUBERT, S. KAZEMI, M. HUBERT, A. CARUGATI, and M. M. MARIANI. 从个性化和虚拟对话代理探索用户感知价值以实现智能家居组合——一种混合方法。*国际信息管理杂志*, 2025, 80: 102850. https://doi.org/10.1016/j.ijinfomgt.2024.102850
[8] H. MANCY, N. E. GHANNAM, A. ABOZEID, and A. I. TALOBA. 用于智能水管理和灾害响应的分散多代理联合和强化学习。*亚历山大工程杂志*, 2025, 126: 8-29. https://doi.org/10.1016/j.aej.2025.04.033
[9] M.-H. YEN, N. MISHRA, W.J. LUO, and C.E. LIN. 一种用于基于 IoE 的智能事物监控系统的新型主动式基于 AI 的代理框架，可用于智能车辆。*计算机、材料和连续体*, 2025, 82(2): 1839-1855. https://doi.org/10.32604/cmc.2025.060903
[10] Y. WANG, S. SU, and Y. WANG. 用于智能工业物联网任务卸载的注意力增强多智能体协作。*物联网*, 2025, 31: 101572. https://doi.org/10.1016/j.iot.2025.101572

[11] L. V. D. NASCIMENTO, and J. P. M. D. OLIVEIRA. 用于智能城市上下文源集成的多代理架构。*未来一代计算机系统*, 2025, 172: 107862. https://doi.org/10.1016/j.future.2025.107862
[12] G. SINGH. 可穿戴物联网 (w-IoT) 人工智能 (AI) 解决方案，实现可持续智能医疗。*国际信息管理数据洞察杂志*, 2025, 5(1): 100291. https://doi.org/10.1016/j.jjimei.2024.100291
[13] B. RAVIKRISHNA, M. E. SENO, M. RAPARTHI, R. R. YELLU, S. ALSUBAI, A. K. DUTTA, A. AZIZ, D. ABDURAKHIMOVA, and J. BHOLA. 智能医疗场景下疾病预防数据集构建的人工智能概率方案。*SLAS 技术*, 2024, 29(4): 100164. https://doi.org/10.1016/j.slast.2024.100164
[14] S. M. SCHUELLER, C. STILES-SHIELDS, and L. YAROSH. 儿童和青少年精神病学的在线治疗和虚拟治疗师。*北美儿童和青少年精神病诊所*, 2017, 26(1): 1-12. https://doi.org/10.1016/j.chc.2016.07.011
[15] Y. P. P. CHEN, C. JOHNSON, P. LALBAKHSH, T. CAELLI, G. DENG, D. TAY, S. ERICKSON, P. BROADBRIDGE, A. E. REFAIE, W. DOUBE, and M. E. MORRIS. 对言语障碍的虚拟言语治疗师进行系统评价。*计算机语音和语言*, 2016, 37: 98-128. https://doi.org/10.1016/j.csl.2015.08.005
[16] A. MARKS, A. GARBATINI, V. P. KIMBERLY HIEFTJE, V. WESER, and C. S. F. FERNANDES. 使用沉浸式虚拟现实空间让青少年和年轻成年癌症患者参与治疗师指导的支持小组：事前研究的方案。*JMIR 研究协议*, 2023, 12: 48761. https://doi.org/10.2196/48761
[17] J. MILLS, and O. DUFFY. 言语和语言治疗师对虚拟现实作为自闭症临床工具的看法：横断面调查。*JMIR 康复和辅助技术*, 2025, 12: 63235. https://doi.org/10.2196/63235
[18] A. MILOFF, P. CARLBRING, W. HAMILTON, G. ANDERSSON, L. REUTERSKIÖLD, and P. LINDNER. 使用虚拟治疗师联盟量表 (VTAS) 衡量自动化治疗时代对具体虚拟治疗师的联盟：开发和心理测量评估。*医学互联网研究杂志*, 2020, 22(3): 16660. https://doi.org/10.2196/16660
[19] K. HAN, and S. KONG. 使用 UTAUT1 探索影响艺术治疗师接受虚拟现实意愿的因素。*心理治疗中的艺术*, 2025, 92: 102260. https://doi.org/10.1016/j.aip.2025.102260
[20] S. G. ROZEVINK, B. MAAS, A. MURGIA, R. M. BONGERS, and C. K. V. D. SLUIS. 治疗师和假肢使用者对为上肢假肢控制训练设计的虚拟现实环境的评估。*手部治疗杂志*, 2025. https://doi.org/10.1016/j.jht.2025.01.004
[21] T. AHMED, E. L.-G. M. R. ISLAM, I. WANG, and M. RAHMAN. 使用 SREx 进行机器人辅助康复：一种用于减轻治疗师身体压力的智能机器人外骨骼。*物理医学与康复档案*, 2022, 103(12): 184-185. https://doi.org/10.1016/j.apmr.2022.08.936
[22] T. H. NGUYEN, M. ELIOS, A. M. FITZPATRICK, S. M. NYENHUIS, and W. PHIPATANAKUL. 优化儿童哮喘管理：AIR 和 SMART 疗法的实施和评估。*《过敏与临床免疫学杂志：实践》*, 2025. https://doi.org/10.1016/j.jaip.2025.05.004
[23] W. BAE, N. JUNG, Y. JUNG, and K. KAM. 使用 AI-motion SW 分析人工智能技术对手部运动受限或不受限制的人的手部运动进行分类的适用性：一项观察性研究。*雪绒花应用科学与技术*, 2025, 9(3): 2930–2940. https://doi.org/10.55214/25768484.v9i3.5889
[24] T. V. P. CISNEROS, A. BRONS, B. KRÖSE, B. SCHOUTEN, and G. LUDDEN. 脑瘫儿童手部治疗中的嬉戏和新技术：范围审查。*JMIR 严肃游戏*, 2023, 11(1): 44904. https://doi.org/10.2196/44904
[25] G. BURDEA, N. KIM, K. POLISTICO, A. KADARU, N. GRAMPUROHIT, D. ROLL, and F. DAMIANI. 用于中风后人工智能增强远程康复的辅助游戏控制器。*辅助技术*, 2019, 33(3): 117–128. https://doi.org/10.1080/10400435.2019.1593260
[26] D. N. SOUMIS, and N. D. TSELIKAS. 基于网络的手部康复评估平台。*大数据与认知计算*, 2025, 9(3): 52. https://doi.org/10.3390/bdcc9030052
[27] P. MANIAR, G. VEDPATHAK, and S. FRANCIS. VTherapy：基于深度学习的手部康复系统。在 *第二届医疗保健人工智能国际会议*，印度瓦尔达教育与工业（IDICAIEI）, 2024. https://doi.org/10.1109/IDICAIEI61867.2024.10842730
[28] J. GONZALEZ, and W. YU. 使用 LSTM 神经网络进行非线性系统建模。 *IFAC 在线论文*, 2018, 51(13): 485-489. https://doi.org/10.1016/j.ifacol.2018.07.326
[29] W. SIMOES, L. REIS, C. ARAUJO, and J. M. JR. 使用 MediaPipe 对物理治疗练习的 2D 姿势估计进行准确度评估。*Procedia 计算机科学*, 2024, 251: 446-453. https://doi.org/10.1016/j.procs.2024.11.132
[30] A. LATRECHE, R. KELAIAIA, A. CHEMORI, and A. KERBOUA. 基于 MediaPipe 的某些人体康复动作测量系统的可靠性和有效性分析。*测量*, 2023, 214: 112826. https://doi.org/10.1016/j.measurement.2023.112826
[31] S. RANI, S. MASOOD, and D. R. RIZVI. 基于 Bi-LSTM 的机器人中风康复治疗期间补偿检测方法。*Procedia 计算机科学*, 2025, 258: 3145-3154. https://doi.org/10.1016/j.procs.2025.04.572
[32] A. WOLFF, G. WEINSTOCK-ZLOTNICK, and J. GORDON. SSc 管理——面对面预约和远程治疗 (SMART)：管理慢性手部疾病的框架。*手部治疗杂志*, 2014, 27(2): 143-151. https://doi.org/10.1016/j.jht.2013.11.004

Word count (excluding references): 7,335 words.

Peer-review record:

- *Fast-track status:* Not fast-tracked
- *First-round reviews received:* 3 reports
- *Revision cycles completed:* 3 rounds
- *Final version submitted:* July 2, 2025