

AI-Driven Stroke Rehabilitation Systems and Assessment: A Systematic Review

Sejuti Rahman¹, Sujan Sarker, *Student Member, IEEE*, A. K. M. Nadimul Haque¹,
Monisha Mushtary Uttsha¹, Md Fokhrul Islam¹, and Swakshar Deb

Abstract—Post-stroke therapy restores lost skills. Traditionally, patients are supported by skilled therapists who monitor their progress and evaluate the program’s effectiveness. Due to a shortage of qualified therapists, rehabilitation facilities are both expensive and inadequate. Furthermore, evaluations may be subjective and prone to errors. These limitations motivate the researchers to devise automated systems with minimal human intervention, therapist-like assessment, and broader outreach. This article reviews seminal works from 2013 onwards, qualitatively and quantitatively adapting the PRISMA approach to examine the potential of robot-assisted, virtual reality-based rehabilitation and automated assessments through data-driven learning. Extensive experimentation on KIMORE and UI-PRMD datasets reveal high agreement between automated methods and therapists. Our investigation shows that deep learning with spatio-temporal skeleton data and dynamic attention outperforms others, with an RMSE as low as 0.55. Fully automated rehabilitation is still in development, but, being an active research topic, it could hasten objective assessment and improve outreach.

Index Terms—Automated stroke rehabilitation, VR-aided rehabilitation, robot-assisted rehabilitation, data-driven assessment, deep learning.

I. INTRODUCTION

STROKE is a common medical condition in which blood supply to the brain is cut off, resulting in cell death. It is a leading cause of disability worldwide [1]. A stroke may have short or long-term consequences, as around 35% of patients have deterioration in their cognitive and physical abilities [2]. However, many stroke patients can repair and re-learn motor functions during the therapeutic time, according to study [3], demonstrating the efficacy of rehabilitation therapy. Conventionally, stroke rehabilitation begins when a physician determines the severity of the stroke. This rehabilitation process can be of two types: inpatient rehabilitation, which takes place in a hospital, and outpatient rehabilitation, which takes place in the patient’s home. Inpatient rehabilitation treatments often provide superior outcomes [4] as a consequence of the

therapists’ continual supervision. On the contrary, outpatients seldom have therapists there to help them through the recovery process. Thus, the patients doing the exercises are not monitored or assessed except during the follow up assessments. This leads to a possibility that the patients are not doing the exercises properly, resulting in subpar outcomes.

Despite the benefits of inpatient rehabilitation, treating every patient in the hospital is not always practical. One of the primary reasons for this is that it needs a significant amount of physicians. There are around 300 trained therapists for every million individuals in affluent nations such as the United States or Australia [5]. This results in around 65% of persons in the United States obtaining rehabilitation treatments [4]. This is exacerbated in developing nations, where about 70% of all stroke cases occur [6]. Given the high rehabilitation demands in these nations, there are less than ten qualified therapists per million persons [7]. In the United States, the average lifetime cost of inpatient rehabilitation and follow-up care is around 140,048 dollars per person [8]. Another issue that emerges in both inpatient and outpatient settings is that exercise evaluations, when they occur, are conducted by humans. This might result in subjective and biased judgments, and ultimately hinder the progress. While the expense of rehabilitation can not be eliminated instantly, the paucity of qualified therapists, the subjective nature of evaluation, and even the shortcomings of at-home therapy may all be addressed by automating the stroke recovery process.

Automated stroke therapy systems can be classified into two categories: robot-assisted and virtual reality-based. In a robot-assisted system, the patient is aided in doing the exercises by a therapy robot or an exoskeleton. On the contrary, in VR-assisted systems, patients are immersed in a virtual environment, often a game, that assists them in doing the exercises. In all cases, the overall process can be broken down into three smaller tasks: assisting with exercises and subsequent monitoring to gather data, evaluating the data, and producing correct conclusions about the exercise’s quality. Each of these three jobs is automatable. However, our review of the literature reveals that automated control of the peripheral devices used in stroke rehabilitation systems is almost never used. This is possibly due to the hazards associated with such devices such as device malfunctions, and abrupt behaviours in the presence of the vulnerable patients. However, data collecting and analysis are automated in many works, provided that the peripherals are correctly installed. Additionally, therapists still have a part to play in automated systems regarding initial

Manuscript received 28 June 2022; revised 7 October 2022; accepted 25 October 2022. Date of publication 3 November 2022; date of current version 30 January 2023. This work was supported in part by the Innovation Fund of ICT Division, Ministry of Posts, Telecommunications and Information Technology, Government of Bangladesh; and in part by the Centennial Research Grant, University of Dhaka. (*Corresponding author: Sejuti Rahman.*)

The authors are with the Department of Robotics and Mechatronics Engineering, University of Dhaka, Dhaka 1000, Bangladesh (e-mail: sejuti.rahman@du.ac.bd).

Digital Object Identifier 10.1109/TNSRE.2022.3219085

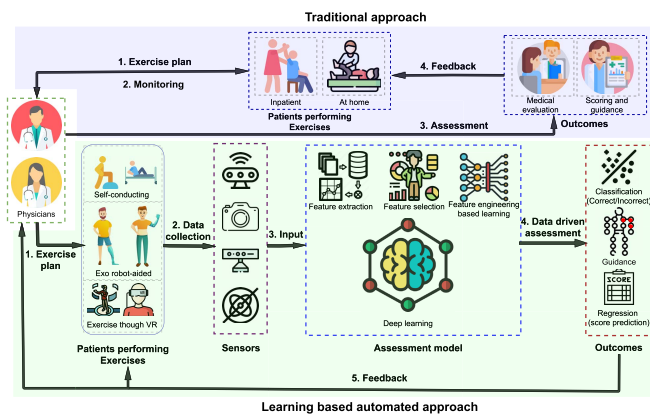


Fig. 1. A comparative overview between traditional vs automated post-stroke rehabilitation systems.

diagnosis, stroke severity analysis and planning the rehabilitation program. A comparative overview between traditional vs automated post-stroke rehabilitation systems is drawn in Fig. 1. The diagram above depicts the traditional process, which need expert assistance at each stage. The image below depicts the automated procedure, which still involves therapists in the rehabilitation planning. However, the next steps only require minimum human intervention. The second step (data collection) entails gathering information from a variety of sensors (Kinect v2 [9], Gyroscope [10], IR cameras [11], Vicon [12]). Step three involves feeding this input data into an assessment model, which then gives performance evaluation and recommendations in step four. This is then forwarded to both the doctor and the patient as feedback so that the rehabilitation plan, as illustrated in step five, may be revisited as well as enables patients to comprehend their performance and make necessary adjustments.

A. Related Surveys

This section reviews surveys on automated rehabilitation systems, broadly classifying them into two subcategories.

1) *Robot-Assisted Stroke Rehabilitation Systems*: The most common application of robotics in automated rehabilitation systems is actively assisting the patient using an exoskeleton or a robotic manipulator. Frolov et al. [13] investigated the neurophysiological aspects of such devices in rehabilitation, whilst Jarasse et al. [14] reviewed their potential for functional recovery. Additionally, the study in [14] sought to identify flaws in the mechanical designs and different control algorithms for such systems. The authors in [15] explored the role of therapeutic robots such as ARMin [16], the HapticWalker [17], and others in motor function rehabilitation, rehabilitation gyms, and robot-assisted telerehabilitation. However, these studies focused only on robot-assisted systems, leaving the automated evaluation of exercises untouched.

2) *VR-Assisted Stroke Rehabilitation Systems*: Virtual reality-enabled systems emulate real-world experiences in a virtual environment, often via the medium of a game. The patient gets engaged in exercise-like movements while playing games, and their performance is assessed either by the game score or an automated assessment performed by

TABLE I

A COMPARISON BETWEEN OUR STUDY AND RELATED SURVEYS ('-' = 'NOT APPLICABLE', 'QA' = 'QUANTITATIVE ANALYSIS')

Ref.	Focus of Study	No. of Pa-pers	No. of Data sets	QA
[13]	The neurophysiological basis of exoskeleton-based rehabilitation	38	-	-
[14]	Mechanical design of robotic exoskeleton	30	-	-
[15]	Therapy robots in motor impairment rehabilitation	34	-	-
[18]	Gamified stroke rehabilitation	28	-	-
[19]	Design and assessment of VR-based rehabilitation	21	-	-
[20]	Kinect based stroke rehabilitation	55	-	-
[22]	Wearable technologies for gesture and exercise recognition	46	-	-
Ours	VR and robot-assisted rehabilitation and automated assessment	48	8	Yes

machine learning models. Tamayo-Serrano et al. [18] identified 20 distinct features of VR-based systems' quality, including cost, difficulty, and rehabilitation types. However, they did not incorporate any quantitative analysis. Authors in [19] discussed the technical design aspects of VR-based neuromuscular rehabilitation systems and provided a comparative perspective. Webster and Celik [20] conducted a comprehensive assessment of the literature on the use of the Kinect in elderly care and exercise monitoring (e.g., fall detection and risk reduction), as well as exercise games. However, the drawbacks of the surveys regarding robot-assisted systems are persistent in the above surveys as well. Furthermore, none of the aforementioned polls include a quantitative analysis of available solutions.

Apart from the works mentioned above, Langan et al. [21] reviewed stroke rehabilitation technologies via questionnaires, revealing that traditional methods continue to outperform more recent stroke rehabilitation systems that incorporate games, virtual reality, and so on. Besides, authors in [22] covered the wearable devices used to gather patient exercise data, and according to their analysis, wearable gloves give the most accurate measurements of all the devices. It is quite apparent that the surveys of stroke rehabilitation systems are restricted in scope, concentrating on either robot-assisted or VR-assisted systems. Moreover, the techniques for exercise assessment remain unexplored. This has brought out another gap in this field: quantitative analysis of automated exercise assessment. Our research attempts to provide a comprehensive picture of rehabilitation systems by examining works on robot-assisted, virtual reality-based rehabilitation and automated assessments through data-driven learning. Table I provides a summary of the existing survey papers discussed above.

B. Contributions

This is the only review article that we are aware of that has a complete and systematic discussion on automated stroke rehabilitation systems. Rather than focusing on a single aspect or branch of rehabilitation systems, we examine them from a broader perspective to see how state-of-the-art

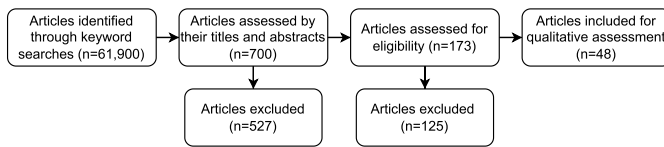


Fig. 2. The article selection process using PRISMA.

technology in automated rehabilitation systems functions. We also shed light on the research challenges and the direction of future research. In addition to the qualitative discussion, we provide a quantitative analysis by evaluating a variety of methodologies using publicly available datasets. In particular, throughout this article, we opt to answer the following research questions:

RQ1: Do automated rehabilitation systems function as efficiently as traditional therapy?

RQ2: Are such efforts properly clinically validated?

RQ3: To what extent does automated diagnosis need human intervention?

RQ4: How closely does an automated method's diagnosis match the physician's assessment?

RQ5: Does deep learning outperform feature engineering-based machine learning algorithms for exercise assessment?

RQ6: Can we move past the existing challenges associated with automated rehabilitation systems and exercise evaluations?

C. Survey Methodology

The works included were chosen using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) technique [23]. As shown in Fig. 2, the selection progressed through four phases: identification of relevant papers; further screening of those based on title and abstract; determining their eligibility against a specific set of criteria; and lastly, inclusion of the chosen works. The works were identified first by searching for “automated stroke rehabilitation systems,” “robot-assisted stroke rehabilitation systems,” “virtual reality-based rehabilitation systems,” and “machine learning-based rehabilitation exercise assessment.” Hence, 61,900 publications relating to stroke rehabilitation systems and evaluation were found. The following criteria apply to the first screening: the article must have been published in a peer-reviewed journal or conference and be deemed relevant to this review based on the title and abstract. Following an initial screening, the paper count was decreased to 700 papers. The next step was eligibility, which considered the overall quality of the article, including qualitative and/or quantitative analysis. This resulted in the final inclusion of only 48 publications on automated stroke rehabilitation systems and assessment, including 24 articles on robot-assisted or VR-based systems, 7 datasets and 17 articles on machine-learning-based exercise evaluation.

The remainder of the paper is structured as follows. In section II, we review the notable works in automated stroke rehabilitation systems. We discuss the work done in automated exercise assessment methods, including available datasets, in section III. Section IV offers a quantitative comparison of

several state-of-the-art assessment methods and a side-by-side comparison of deep learning and feature engineering-based learning approaches. We shed light on some of the existing challenges in the field in section V. Section VI holds a comprehensive discussion of the findings and answers to the research questions presented before. Finally, in section VII, we provide our conclusion and briefly address future prospects.

II. RESEARCH IN AUTOMATED POST STROKE REHABILITATION SYSTEMS

This section reviews notable works in automated stroke rehabilitation, broadly divided into two branches: robot-assisted and virtual reality-assisted stroke rehabilitation systems. Since the exercise evaluation techniques are the same for both branches, it is discussed later in Section III.

A. Robot-Assisted Systems

Robot-assisted stroke rehabilitation solutions include exoskeletons and assistive or demonstrative robots. Although similar systems, e.g., the hybrid assistive limb (HAL) [24], have existed for some time, the cost has always been a concern. However, recent advancements have resulted in a cost reduction and performance on par with, if not better than, traditional treatment [25]. Aprile et al. [26], conducted a study in which three robots assisted 51 patients with shoulder, elbow, and finger flexions while a sensor-based device recorded their movements. To evaluate the patients, cognitive tests (e.g., digit span [27], Oxford cognitive test [28]), motor (FMA and Motricity index [29]) and disability scales (modified Barthel index [30]) were used. These measures improved significantly, demonstrating the effectiveness of robot-assisted treatment. Ou et al. [31] developed a wearable exoskeleton to allow more thumb joint mobility of stroke patients. However, no clinical evaluation of the exoskeleton was conducted.

Pilla et al. [32] suggested a randomized trial method based on a robotic exoskeleton, the NEUROExos Elbow Module (NEEM) to determine the effectiveness of a rehabilitation system. The authors of [33] suggested a clinical study using four distinct robotic setups for upper limb rehabilitation of 190 stroke patients. They reported that the mean FMA score increased from the baseline by 8.50 and 8.57 points, respectively, in the robot-assisted and conventional therapy group. This indicates that, robotic rehabilitation barely outperforms conventional therapy, albeit at a much higher cost. This is consistent with the findings of [34].

A number of studies have examined the use of electromyography (EMG) signals to control robotic systems. For instance, Qian et al. [35] developed an EMG-driven neuromuscular electrical stimulation (NMES) robotic arm for upper limb rehabilitation. They examined the device's influence on the FMA score, the Modified Ashworth Scale (MAS) [36], the ARAT score [37], and the Functional Independence Measurement (FIM) on 24 subacute stroke patients. The results indicated remarkable improvements in all measures rehabilitated by the NMES-robot. Similarly, authors of [38] used EEG and EMG signals for brain intent recognition to develop a brain machine interface (BMI) to enhance the functionality of an exoskeleton.

Baca et al. [39] suggested a wearable, modular robotic system, whereas Lu et al. [40] proposed an EMG-driven exoskeleton hand controlled by the patients' intentions. Apart from [39], which did not have any clinical validations to date, all the others indicate significant improvement in all metrics.

Some studies have sought low-cost alternatives to robotic-assisted rehabilitation systems. The authors of [41] applied their system in a treatment gym, and found the system to be nearly 1.5 times less expensive than traditional one-to-one therapy while still providing better results. Reduced reliance on pricey hardware would be an alternate strategy to increase cost-effectiveness. Using smaller end-effector robotic devices instead of full-scale robots or exoskeletons can lead to less expensive yet effective rehabilitation solutions. Perfect examples are MIT-Manus [42] and REAplan [43]. These, when combined with appropriate software, can form a reliable system. The initial study of [44] had explored the neuro-rehabilitation capabilities of an MIT-Manus based system, finding no negative consequences. This was further proved in the work of Heins et al. [45], that presented ROBiGAME, playable with a REAplan controller for improving stroke patients' motor and cognitive functionality. A feasibility study conducted with two stroke patients observed that patients loved playing the game. However, the impact on stroke patients' rehabilitation has yet to be determined.

B. Virtual Reality Aided Post Stroke Rehabilitation Systems

Virtual reality (VR) technology has experienced tremendous success in recent years, gaining widespread appeal during the previous decade. Rather than having physical rehabilitation sessions, a simulated rehabilitation program can be held. These novel systems are appreciated by the patients as well [46]. It is worth noting that, a few researches found that there is no statistically significant advantage of VR over traditional therapy in terms of overall performance. Adie et al. [47] examined 240 patients with arm weakness who were randomly assigned to daily exercise with a VR device or with simple arm exercises performed at home. Nevertheless, there were no significant variations in mean ARAT scores and other measures between the two groups. Instead, the VR-based rehabilitation would cost the patients an additional £336. Rosiak et al. [48] found similar results with little difference measured in the Vertigo Symptom Scale questionnaire [49]. Schuster-Amf et al. [50] did not find significant improvements in their chosen parameters (Box and Block Test [51], the Chedoke-McMaster Arm and Hand Activity Inventory [52], Stroke Impact Scale (SIS) [53]) either.

In contrast to these, numerous studies revealed that VR-based rehabilitation systems outperform conventional methods. For example, Ho et al. [54] found that VR-based treatment significantly improved NIHSS and modified Ranking Scale (mRS) scores in an investigation with 100 stroke patients. Similarly, Joo et al. [55] saw substantial improvements in Jebsen Taylor hand function test (JTT) [56] and Michigan Hand Outcome Questionnaire (MHQ) [57] scores.

The majority of the works mentioned above have taken a gamified approach, that is, they have built games that can be played in virtual reality and aid in the rehabilitation process. In fact, this approach is predominant in almost all the state-of-the-art works in VR-based rehabilitation. For instance, Warland et al. [58] presented an upper limb rehabilitation system that included both exercise and an apple catching game. Overall, all evaluated parameters improved, with five subjects reporting minor side effects. Elor et al. [59] also presented an immersive VR game for stroke therapy based on the Constraint-Induced Movement Therapy (CIMT) [60]. According to their findings, players expressed an interest in utilizing their stroke-affected arms to enhance game rewards. Compliance with affected arm usage was as high as 78% for easy and medium difficulties, a substantial increase over the 32% compliance rate in traditional CIMT [61].

In contrast to gamified approaches, in which patient rehabilitation is directly linked to in-game scores, a few studies have used external sensors to monitor and assess exercises completed with the assistance of VR systems. For example, Luca et al. [62] performed neurocognitive rehabilitation of patients with the help of the BTs Nirvana system along with an infrared video camera for monitoring and assessment. The results indicated that patients in the VR group progressed much more than those in the control group in cognitive rehabilitation. A similar observation is made in [63], where VR-based balance training improved patients' foot placement performance.

Additionally, a few studies have included internal EEG and EMG signals to aid and assess rehabilitation progress. Vourvopoulos et al. [64] presented a head-mounted brain-computer interface (BCI) for post-stroke rehabilitation based on the REINVENT system [65]. The authors of [66] compared the performance of EMG and EEG signals used as biofeedback in VR-based rehabilitation and found participants performing much better with EMG than with EEG feedback. A summary of the robot-assisted and VR-based works are presented in Table II and Table III, respectively.

Discussion on RQ1 & RQ2

It is evident by the studies mentioned above that the patients cope well with such systems, with initial feasibility assessments indicating high levels of agreement. Such systems are comparable to traditional therapy and often yield greater results, which answers RQ1. In response to RQ2, Tables II and III demonstrate that the majority of the works performed a clinical study. They conducted a number of clinical trials and validations using a variety of patients and therapists, confirming that automated procedures are consistent with the clinical standard.

III. AUTOMATED EXERCISE ASSESSMENT

Automated stroke rehabilitation systems must incorporate some form of assessment to monitor the system's efficacy and ensure optimum recovery. This section summarizes the

TABLE II

OVERVIEW OF EXISTING ROBOT-ASSISTED STROKE REHABILITATION SYSTEMS (Y = YES, N = NO, M = MARGINAL, '-' = NOT STUDIED)

Ref.	Approach/Focus	Is clinical trial done?	No. of subjects (Experimental:Control group)	Evaluation Metric	Does VR outperform conventional therapy?
[32]	Exoskeleton-aided	Y	60(30:30)	FMA, MAS scores	-
[33]	Robot-assisted upper limb rehabilitation	"	190 (91:99)	FMA scores	M
[35]	EMG-driven robot arm for upper limb	"	24(14:10)	FMA, MAS, ARAT and FIM scores	Y
[38]	Exoskeleton with brain machine interface	"	6	FMA scores and movement quality scores	-
[41]	Robot-assisted lower and upper limb exercises	"	20(10:10)	FMA scores, Box and Block, Walk tests	M
[34]	Robot-assisted arm rehabilitation	"	770	ARAT score	N
[40]	Arm rehabilitation with Myo Armband and Leap Motion Controller	"	8	Accuracy	-
[43]	REAPLAN device for upper limb	"	50	ANOVA and PCA	-
[26]	Robot-assisted upper extremity rehabilitation	"	51	Cognitive and motor assessment tests	-
[45]	ROBiGAME with a REAPLAN robotic device	N	2	Likert scale questionnaire	-
[39]	Wearable and modular robotic system	"	-	-	-
[31]	Wearable exoskeleton-aided	"	-	-	-

research on automated exercise evaluation that has been conducted. The majority of the works use some sort of supervised learning, and therefore we begin with a discussion on the datasets available for stroke rehabilitation.

A. Existing Datasets

In this section, we discuss some existing, publicly available datasets that are directly applicable for post stroke rehabilitation exercise assessment.

1) *Selection of Available Datasets*: The PRISMA approach [23] was adopted for selecting datasets. The identification was carried out using the keywords "Stroke rehabilitation exercise dataset" and "Exercise dataset," which resulted in the selection of 12,600 papers. The initial screening took into account the following criteria: dataset availability, suitability for poststroke rehabilitation exercise evaluation, whether the dataset is publicly available or not and whether the datasets are meant for action recognition but can be used in this context. After the initial screening, the number of articles was reduced to 200. Following that, the datasets' eligibility was verified, resulting in a further reduction to 25. The primary criterion was that the dataset be properly annotated and that the data included be of high quality. This reduced the number of datasets presented to only seven.

2) *Overview of Selected Datasets*: Here, we briefly describe the datasets deemed to be related to the field of automated stroke rehabilitation systems and assessment. Table IV

presents a summarized description of the publicly available datasets.

a) *IntelliRehabDS (IRDS) dataset [67]*: The dataset captured the 3D data of nine gesture movements performed by 29 subjects, with 15 being patients and the rest, healthy. Two separate professionals annotated the type of gesture, position of the subjects, as well as the correctness of the movements with an agreement level of 88%. The patients were given the choice to perform the movement either sitting or standing while the healthy subjects perform the both. Even though the dataset consists of both correct and incorrect movements with appropriate labels, the ratio of them is highly unbalanced.

b) *Quality of movement assessment for rehabilitation dataset (QMAR) [68]*: It is worth mentioning that the dataset is comprised of movement data from 38 healthy participants who were simulated to have Parkinson's disease, a stroke, or a limp. A physiotherapist trained the individuals to do two movements: walking and sitting up and down, and their performance was graded in three different ranges. The dataset contains RGB, depth, and skeleton data pertaining to the individuals' motions, with six views accessible for the RGB data and two views available for the depth and skeleton data.

c) *The multi-modal exercise (MEx) dataset [69]*: The dataset contains 6262 occurrences of seven exercises carried out by thirty people. 47% of individuals were between the ages of 18 and 24, while the remainder were between the ages of 24 and 54. At the start of each exercise, volunteers were instructed to complete it for a maximum of 60 seconds while

TABLE III
OVERVIEW OF EXISTING VR-BASED STROKE REHABILITATION SYSTEMS (Y = YES, N = NO, M = MARGINAL, '-' = NOT STUDIED)

Ref.	Approach/Focus	Is clinical trial done?	No. of subjects (Experimental:Control group)	Evaluation Metric	Does VR outperform conventional therapy?
[64]	Brain computer interface (BCI) for rehabilitation	Y	1	FMA scores, Stroke impact scale, Simulator Sickness questionnaire	-
[58]	Nintendo Wii based personalized rehabilitation system	"	12	11 point Likert Scale, FMA score, ABIL-HAND, ARAT score	-
[47]	Nintendo Wii Sports™ based rehabilitation	"	209 (101:108)	ARAT scores	M
[50]	VR vs. conventional therapy in upper limb rehabilitation	"	54 (22:32)	Box and Block Test, Chedoke-McMaster Arm and Hand Activity Inventory, Stroke Impact Scale	N
[54]	Upper limb rehabilitation	"	200 (100:100)	NIHSS score, modified Ranking Scale (mRS)	Y
[55]	Right wrist therapy	"	57 (28:29)	Jebsen Taylor hand function test, Michigan Hand Outcome Questionnaire	"
[62]	Rehabilitation using BTs-Nirvana	"	12 (6:6)	Montreal Cognitive Assessment, Trunk Control Test	"
[63]	Balance improvement therapy	"	10 (5:5)	Wilcoxon Signed Rank Test	"
[46]	Gamified upper extremity rehabilitation	"	16 (8:8)	Feasibility study through participants' reports	Y
[66]	EMG-driven REINVENT architecture for therapy	"	4	Subjective scores, Biofeedback signal comparisons	-
[48]	Vestibular rehabilitation	"	50 (25:25)	Vertigo Symptom Scale questionnaire	Y
[59]	HTC Vive based gamified rehabilitation	"	9	Interviews of subjects, caretakers and therapists	-

their movements were recorded using two accelerometers, a pressure mat, and a depth camera.

d) *KIMORE dataset* [70]: The subjects were divided into two groups: the control group (CG) and the pain and postural disorders (GPP) group with chronic motor disabilities. The CG group is further divided into two subgroups: expert physiotherapists and non-experts. Clinicians chose the following five exercises: upper limb movement stretching the trunk muscles, trunk movement in each of the three planes, and lower limb movement. Experts evaluated each exercise by watching the videos and responding to a ten-item Likert questionnaire [71]. Three scores were calculated based on the responses: the clinical total score, the clinical primary outcome score, and the clinical control factors score.

e) *UI-PRMD dataset* [76]: The dataset contains joint angle and position measurements from ten healthy people performing ten exercises. A Vicon optical tracker and a Kinect camera were used for capturing the exercise motions from 22 different joints. Ten episodes of the ten exercises were executed suboptimally by the individuals and are included as instances of incorrect movements for the test set. The UI-PRMD dataset only has correct/incorrect (binary) annotation because it was originally designed for classification tasks. An annotation scheme based on a gaussian mixture model was later proposed by Liao et al. [77] for producing assessment scores. This model determines the deviation from the ideal movement patterns, which are obtained from healthy patients

executing the exercises to their fullest potential. A monotonic scoring function is used to further transform this deviation into the 0–1 range. This scoring system may predict movement patterns with a satisfactory level of justification, according to experimental results from [77] and [87], with the reference movement receiving a higher score and the incorrect movement receiving relatively low scoring.

f) *AHA-3D* [83]: The dataset contains 79 skeleton videos of exercises by 21 subjects. The skeletal data was collected using Kinect v2 3D cameras in conjunction with RGB cameras, and raw data were labeled using a custom-built GUI. The subjects were instructed to perform four exercises: a 30-second chair stand to assess lower-body strength; an eight-foot up and go to assess fall risk; a two-minute step test to determine functional fitness, and a unipedal stance to check static balance.

g) *Toronto rehab stroke pose (TRSP) dataset* [84]: This dataset includes motion data from subjects performing stroke rehabilitation exercises. The stroke survivors had experienced a subacute or chronic stroke with disability of their upper limbs. A 2 DOF haptic robot was used to assist in shoulder and elbow movement rehabilitation. In addition to the scripted motions used with stroke patients, healthy participants were instructed to do motions imitating common post-stroke compensatory movements. The motions were classified and labeled by two specialists as following: no compensation, leaning forward, shoulder elevation, and trunk rotation.

TABLE IV
OVERVIEW OF PUBLICLY AVAILABLE DATASETS

Ref.	Date of Publication (DD/MM/YYYY)	Objective	Description (No. of Subjects/Experimental:Control group)/No. of exercises	Sensors used/Type of Data	Annotated by	Related works
IRDS [67]	22/03/2021	Assessment	29 (15:14) / 9 gestures	Microsoft Kinect One / Joint position data	2 annotators	-
QMAR [68]	15/01/2021	"	38 (38:0) / 2 movements simulating 2 different ailments	Primesense camera / Depth and skeletal	Not mentioned	[68]
MEX [69]	20/09/2019	Recognition, assessment	30 (0:30) / 7 exercises	Pressure mat and a depth camera / Co-ordinate, pressure and depth	"	[72]–[74]
KIMORE [70]	12/09/2018	Assessment	78 (34:44) / 5 exercises	RGB-D sensors / Depth, joint position	Expert therapists	[68], [75]
UI-PRMD [76]	11/01/2018	"	10 / 10 exercises	Vicon optical tracker and Kinect camera / Joint position data	Not mentioned	[77]–[82]
AHA-3D [83]	03/09/2018	Recognition	21 / 4 tasks	Kinect v2 / Skeletal videos	Manually with GUI	-
TRSP [84]	15/08/2017	"	19 (9:10)	Kinect v2 / Joint positions	2 experts	[85], [86]

B. Works in Automated Exercise Assessment

Automated exercise assessment can be regarded as a classification task, categorizing a movement into correct or incorrect, or a regression task, predicting the score of a movement. We continue our discussion by grouping the works in automated exercise assessment based on the task type and its significance.

1) *Exercise Correctness Classification*: Exercises can be classified according to the accuracy with which they are performed. According to the literature, feature engineering-based algorithms have been the most frequently used. For instance, in [88], the authors used K nearest neighbor and SVM classifiers to identify compensatory motions in the pressure distribution, achieving F1 scores as high as 0.993. Jung et al. [89] on the other hand used model trees [90] and found modest results with an F-measure of 79.29 percent and a ROC of 0.91.

Lee, in [91], used a hybrid approach that combines a rule-based knowledge model and a predictive model for classifying the quality of motion as 0, 1, or 2. The findings indicated a good agreement level with the therapists' assessments. On this basis, Lee et al. [92] investigated several such hybrid models using a variety of classifiers, including Neural Networks (NNs), SVMs, and others and discovered that NNs produce an effective result. This was further demonstrated in [93], which combined reinforcement learning with a variety of classifiers. In [94], the authors developed an ensemble learning model composed of 18 classifiers, each trained on a random subspace. Using six categories, they found 92% accuracy for Brunnstrom and 82 percent for FMA scoring systems.

Another branch in the literature investigated the potential of deep learning for exercise quality classification. Zhi et al. [85] investigated the classification of compensatory motions in rehabilitation using both SVM and RNN classifiers where RNN did not perform as expected. Kaku et al. [95] also achieved unsatisfactory results with their CNN architecture paired with embedding modules, with an average accuracy of 70%. However, the work of [96] obtained great results, even in semi and uncontrolled environments. Zhu et al. [97] was also proven successful with their suggested multipath CNN, which was composed of a dynamic convolutional network

called D-CNN and a state transition probability CNN called S-CNN, claiming a test accuracy level greater than 90%.

2) *Exercise Quality Score Prediction*: Rather than anticipating discrete class labels as in classification approaches, the work in this section attempts to assign a continuous value as an assessment score compared to that of a professional therapist. Typically, prior research in this field employs a distance function to assess the quality of performed and prescribed exercise [99], [100]. These approaches necessitate multiple pre-processing phases, impeding the system's end-to-end processing. In fact, hand-crafted feature based works are still prominent in the literature. The works of [101] and [89] found great results with their selection of features. Lee et al. [102] calculated the range of motion, their smoothness, and the occurrence of correct and erroneous movements. These features also worked well in score prediction, achieving a high level of agreement with the therapists. In another work, Liao et al. [77] presented a Spatio-temporal network capable of evaluating an exercise. They enhanced performance by combining temporal pyramids, multi-branch convolution, and recurrent layers. They used convolutions on tensors of joint data to disrupt the natural graph structure's fine spatial organization within the human body. Their approach yielded an average absolute deviation of only 0.02527.

3) *Data Input Length*: Input length of the sensor data plays a vital role in assessing the exercises. In many cases, the patient might perform exercises at different speeds and the assessment score should not vary if they are performed correctly. However, most approaches keep the length of the exercise video fixed while training models, causing useful information to be left out. Some works such as [96] and [95] attempt to address this by choosing an input length that provides sufficient to capture movement data of all patients. But if a test sample contains useful information exceeding the predefined length, the model might fail to provide consistent results. One technique for dealing with this is dynamic time warping [104], which can handle temporal sequences of different lengths. This has been a widely used technique of past assessment tasks [100], [103], [105], [106] and even in recent works [107], [108]. A better solution was proposed by Zhu et al. [97] that trained a separate CNN network, D-CNN, to deal with dynamic input lengths. Recently, Deb et al. [87] proposed graph convolution based architecture for assessing the rehabilitation exercise to handle

TABLE V

EXISTING REHABILITATION ASSESSMENT WORKS (SDP = STRUCTURAL DATA PRESERVATION, CC = CORRECTNESS CLASSIFICATION, SP = SCORE PREDICTION, FE = FEATURE ENGINEERED, DM = DEEP MODEL, ACC = ACCURACY, MAD = MEAN ABSOLUTE DEVIATION)

Study	System Task	Learning Type	Algorithm	Fixed Input length?	Region of Interest	SDP	No. of Subjects (Stroke:Healthy)	Performance
[88]	CC	FE	KNN, SVM	Y	Upper extremity	N	8 (8:0)	F1 scores: 0.993 (Binary), 0.981 (Multi-class)
[89]	CC, SP	"	Model trees	"	"	"	5 (5:0)	F1 score: 79.29%, RMSE: 0.32
[91]	CC	"	Multimodal Modal	"	Elbow joint	"	26 (15:11)	F1 score: 0.8565
[92]	"	"	Hybrid model(ML + Rule based)	"	"	"	26 (15:11)	F1 score: 0.8279
[93]	"	"	RL+Classifier	"	"	"	26 (15:11)	F1 score: 0.8119 (RL+NN)
[94]	"	"	Ensemble subspace discriminant analysis	"	Hands	"	50 exercise samples ¹	Acc: 92%
[85]	"	"	SVM, RNN	"	Upper limbs	"	19 (9:10)	F1 scores: 0.82 (SVM), 0.27 (RNN)
[95]	"	"	CNN	"	Upper extremity	"	48(48:0)	Acc: 70%
[96]	"	DM	Rehab-Net (CNN)	"	Arm	"	D1 - 4 (4:0), D2 - 10 (10:0) ²	Acc: 97.89% (D1), 88.87% (D2)
[97]	"	"	Multipath CNN	N	Upper extremity	"	WHARF [98], 49 (0:49) ²	Acc: 90.63%
[99]	"	FE	DTW Distances with KNN	N	Upper limb	N	29 (21:8)	Acc. 82.1
[100]	"	"	DTW distances	"	Upper extremity	"	5 (0:5)	Acc: 91.9% (Posture), 93.75% (Trajectory)
[101]	SP	FE	SVR	Y	Shoulder and elbow	"	24 (24:0)	RMSE: 2.1273
[77]	"	"	CNN, LSTM, HNN	"	Full body	Y	UI-PRMD [76]	Abs. deviation: 0.00656 (CNN)
[102]	CC, SP	"	Binary classifier with threshold modelling	"	Elbow joint	N	26 (15:11)	F1 score: 0.8436
[87]	SP	DM	STGCN	N	Full body	Y	KIMORE [70] and UI-PRMD [76]	MAD: 0.57 (KIMORE), 0.011 (UI-PRMD)
[103]	CC	FE	DTW distances	N	Full body	"	5 (0:5)	Acc: 93.46%

¹ Population details not given ² Multiple datasets used

variable length and preserve local connectivity between each joint.

Table V presents a summarized description of existing literature. One key observation is that only [97] and [87] can deal with varied length inputs. Another issue is that most works adhere to the upper body rehabilitation exercise. It would be better to incorporate lower body rehabilitation exercise data to make the algorithms more robust. Literature suggests that preserving the topological structure of human skeleton has become the trend in action recognition tasks for its great performance [109], [110], [111]. In this line of works, recently Deb et al. introduced graph convolution with a dynamic attention module to assess rehabilitation exercises. Their results significantly outperform the current research, demonstrating the efficacy of structure-aware learning.

Discussion on RQ3 & RQ4

Our analysis here is tailored towards the research questions 3 and 4. In the 3rd research question about therapist involvements, all of the works report the involvement of therapists in one way or another. Thus, these systems do not eliminate the need for therapists, but rather assist them and alleviate some of the dependency on them. The consequent research question is directly addressed by the authors of [91], [92], [102] who directly report high agreement levels among therapist and automated assessment.

IV. QUANTITATIVE ANALYSIS

A qualitative analysis of the existing literature is presented in Section III, where the datasets, number of subjects, and evaluation protocols differ. This non-uniformity prevents us from comparing the results of different methods on the same scale. Therefore, we conduct a quantitative study in which automated assessment methods are evaluated on two datasets: UI-PRMD [76] and KIMORE [70], using following performance metrics: *Mean Absolute Deviation (MAD)*, *Mean Absolute Percentage Error (MAPE)*, and *Root Mean Squared Error (RMSE)* [77], [87]. It should be noted that the results reported in this section are generated in our own experimental setup rather than being directly adapted from the research we compared.

To draw the comparative analysis, we chose seven deep learning methods, including [87] and [77], which were originally proposed for rehabilitative exercise assessment. Since there is a limited corpus of work on exercise assessment in the extant literature, the remaining five models are drawn from human action recognition. Existing work on action recognition can be broadly classified as: CNN, LSTM, and graph convolution-based approaches. Therefore, we choose [115] as CNN, [114], [117] as LSTM, and [87], [109], [112], [113] as graph-based regressors. For the feature engineering-based learning, we apply four regressors: K-Nearest Neighbour (KNN), Random Forest (RF), Support Vector Machine (SVM), Neural Network (NN), on features generated by [116] and [102].

TABLE VI

ANALYSIS OF RESULTS BETWEEN DEEP LEARNING AND FEATURE ENGINEERING-BASED LEARNING APPROACHES ON KIMORE DATASET. THE BOTTOM ROW INDICATES AVERAGE VALUE ACROSS ALL FIVE EXERCISES

Exercise No.	Metric	Deep Learning							Feature Engineering-based Learning							
		Deb et al. [87]	Song et al. [112]	Zhang et al. [113]	Liao et al. [77]	Yan et al. [109]	Du et al. [114]	Li et al. [115]	Capecci et al. [116]				Lee et al. [102]			
									K-NN	RF	SVM	NN	K-NN	RF	SVM	NN
1	MAD	0.79	0.97	1.75	1.14	0.88	1.27	1.37	1.38	4.77	3.12	3.08	5.71	6.96	5.47	5.86
	RMSE	2.02	2.16	2.91	2.53	2.01	2.44	2.34	2.33	5.93	4.34	4.35	7.41	8.90	7.17	7.97
	MAPE	1.92	2.61	5.05	2.58	2.33	3.22	3.49	3.90	21.34	23.82	8.51	19.29	24.05	19.17	18.68
2	MAD	0.77	1.28	3.13	1.52	2.09	2.19	1.87	3.19	4.82	2.83	3.97	7.04	8.29	6.28	7.57
	RMSE	2.12	3.34	4.14	3.73	3.26	4.29	2.82	4.22	5.71	3.68	5.22	8.99	9.64	7.83	9.56
	MAPE	1.27	3.29	10.43	3.97	6.13	6.00	5.29	8.65	21.54	23.53	10.50	27.70	30.96	24.27	25.02
3	MAD	0.36	1.11	1.73	0.84	0.60	1.12	1.45	3.38	5.08	3.94	4.53	5.57	6.50	5.86	7.32
	RMSE	0.55	1.93	2.61	1.56	0.79	1.92	2.00	5.23	6.42	5.40	6.55	7.40	8.31	7.95	9.24
	MAPE	0.72	2.96	5.77	2.02	1.72	3.42	4.18	9.67	19.23	20.89	12.26	18.13	21.48	20.24	21.98
4	MAD	0.34	0.71	1.20	0.46	0.84	0.88	0.67	2.41	4.86	3.34	4.76	5.33	8.18	6.59	7.54
	RMSE	0.64	2.01	1.83	0.79	1.33	1.67	1.07	3.73	5.99	4.52	6.94	7.53	9.47	8.13	10.08
	MAPE	0.84	2.15	3.90	2.33	2.32	2.58	1.97	5.95	19.33	21.36	12.58	18.14	27.08	21.52	24.19
5	MAD	0.62	1.53	1.85	0.84	1.21	1.86	1.66	2.71	4.80	3.04	2.34	3.71	5.57	4.37	5.02
	RMSE	1.18	3.19	2.91	1.91	1.95	3.15	2.57	4.28	6.07	4.91	3.15	5.53	7.71	6.26	7.24
	MAPE	1.59	4.95	6.53	2.31	3.80	5.62	5.75	6.84	16.74	18.04	5.41	12.39	17.90	14.14	16.50
avg	MAD	0.57	1.12	1.92	0.96	1.12	1.46	1.38	2.89	5.82	4.13	4.35	6.61	8.49	6.81	7.83
	RMSE	1.71	2.96	3.46	2.61	2.27	3.18	2.63	4.42	7.21	5.44	6.11	8.85	10.58	8.91	10.41
	MAPE	1.65	3.71	7.35	3.16	2.72	4.81	4.83	7.78	23.91	26.29	11.55	22.98	29.10	23.70	25.01

The experimental results on the KIMORE dataset are presented in Table VI. We observe that, Deb et al. [87] (avg. MAD 0.576) outperforms other methods in all the studied metrics. This is articulated by the fact that they extended the Spatio-temporal Graph Convolution Network (STGCN) with a dynamic self-attention mechanism to extract discriminative features from the structural information of the human skeleton. Furthermore, because each exercise focuses on the movement of a specific set of body joints, this self-attention mechanism makes it easier for the network to include the role of joints in varied exercises. In the squatting exercise (Ex 5), for example, the ankle, knee, spine, and shoulder movement are critical, whereas in the lifting arms exercise (Ex 1), the elbow, spine, thumb, and wrist joint are more important than the rest of the joints. Furthermore, they employ LSTM instead of global average pooling to capture subtle sequential dependencies residing in consecutive frames and to extract discriminative temporal features from variable-length exercise data. The approach by Liao et al. [77] (avg. MAD 0.960) came in second best that adapted a 2D convolution based method. However, 2D convolution treats the skeleton sequences in a grid like fashion that unable to utilise the subtle information contained in spatial characteristics. Yan et al. [109] (avg. MAD 1.124), used an STGCN that overlooked the sequential aspect of the spatio-temporal features because of successive global average pooling. Similarly, Song et al. [112] (avg. MAD 1.120) proposed an STGCN utilizing multi-stream information with joint attention mechanism. Furthermore, Li et al. [115] (avg. MAD 1.384) proposed a hierarchical method as well and aggregated features from point level to global co-occurrence features. Multiple box filters are used in this hierarchical convolution-based model to extract skeleton features, however the spatial relationship between adjacent joints is missed, making it unable to capture topological information. Finally, Du et al. [114] (avg. MAD 1.464) employed a hierarchical RNN based approach. However, the RNN based approaches

only consider the temporal information but overlooks the spatial information. Among the deep learning models, the lowest agreement level with the therapists' assessments is acquired by the approach in [113] (avg. MAD 1.920) as their model does not capture significant sequential connections between subsequent frames due to many spatial and temporal maxpooling layers.

Capacci et al. [116] and Lee et al. [91] adopted feature engineering-based learning on KIMORE dataset to predict the movement correctness score. However, the use of several preprocessing stages restricts the model's ability to capture meaningful information from the movement data.

We conducted experiments on the UI-PRMD dataset in Table VII. Here we found that graph convolution based approaches, i.e., Deb et al. [87], Song et al. [112], Zhang et al. [113] consistently outperform the rest of the approaches. However, interestingly, Lee et al. [91], and Capecci et al. [116], although being feature engineering-based approaches, perform better than the rest of the deep learning methods. The reason for this is that the UI-PRMD dataset contains a homogeneous distribution of patients (i.e., ten healthy patients performing 10 exercises correctly and incorrectly) with a small number of training samples. As a result, movement patterns became less varied. As a consequence, unlike graph convolution-based models, CNNs are unable to discern subtle differences associated with each movement pattern due to disregarding structural information.

A. Results Using RGB Camera

With the advancement of technology, depth cameras like the Vicon, Kinectv1 [9], and Kinectv2 [9] are now more affordable for use in healthcare facilities. However, rather than solely relying on the advanced RGBD sensors, in this section, we analyze the potential of using RGB cameras,

TABLE VII
PERFORMANCE OF DEEP LEARNING VS. FEATURE ENGINEERING-BASED APPROACHES ON UI-PRMD DATASET USING MAD

Exercise No.	Deep Learning							Feature Engineering-based Learning							
	Deb et al. [87]	Song et al. [112]	Zhang et al. [113]	Liao et al. [77]	Li et al. [115]	Shahroudy et al. [117]	Du et al. [114]	Capecci et al. [116]				Lee et al. [102]			
								K-NN	RF	SVM	NN	K-NN	RF	SVM	NN
1	0.009	0.011	0.022	0.011	0.011	0.018	0.030	0.042	0.040	0.041	0.039	0.017	0.027	0.020	0.020
2	0.006	0.006	0.008	0.028	0.029	0.044	0.077	0.008	0.008	0.009	0.018	0.007	0.010	0.007	0.007
3	0.013	0.010	0.016	0.039	0.056	0.081	0.137	0.015	0.018	0.016	0.023	0.015	0.013	0.012	0.018
4	0.006	0.014	0.016	0.012	0.014	0.024	0.036	0.018	0.021	0.021	0.025	0.009	0.010	0.008	0.009
5	0.008	0.013	0.008	0.019	0.017	0.032	0.064	0.019	0.019	0.018	0.019	0.013	0.012	0.010	0.012
6	0.006	0.009	0.008	0.018	0.019	0.034	0.047	0.014	0.014	0.014	0.014	0.007	0.009	0.007	0.007
7	0.011	0.017	0.021	0.038	0.027	0.049	0.193	0.016	0.018	0.019	0.020	0.016	0.017	0.017	0.020
8	0.016	0.017	0.025	0.023	0.024	0.050	0.073	0.014	0.020	0.016	0.022	0.024	0.025	0.024	0.030
9	0.008	0.008	0.027	0.023	0.027	0.043	0.065	0.016	0.016	0.017	0.018	0.012	0.012	0.013	0.014
10	0.031	0.038	0.066	0.041	0.046	0.077	0.160	0.076	0.091	0.085	0.095	0.049	0.051	0.057	0.067
avg	0.011	0.014	0.021	0.025	0.027	0.045	0.088	0.024	0.027	0.026	0.030	0.017	0.019	0.018	0.021

a more economical solution for the assessment task. Instead of only relying on the more sophisticated RGBD sensors, in this part we emphasize the more practical approach of using RGB cameras for the assessment task. Since RGB cameras are unable to provide skeleton information, we select two 3D pose estimation algorithms: videopose3D [118] and blazepose [119], to detect joint information automatically from RGB information. Finally, we compare the performance with RGBD sensors, i.e., Kinectv2 [9].

In Table VIII, we analyze the performances of the studied approaches on RGB videos of patients performing exercises as provided by the KIMORE dataset. To extract the 3D poses from videos, we used two pose estimation algorithms: BlazePose [119], (trained on MS Coco [120]) and VideoPose3D [118] (trained on Human3.6M [121]). We also compare the results against the pose data extracted by Microsoft Kinectv2 [9]. From the results, we observe that the Microsoft Kinectv2 sensor surpasses the other two pose estimation methods based on RGB cameras because it uses RGBD information to detect human poses. In addition, GCN based algorithms [109] perform well on BlazePose [119] because they can exploit spatial information better than CNN based approaches [77], [114], [115]. In the case of Deb et al. [87], we found that BlazePose outperforms VideoPose3D since it models the human skeleton with more joints (38 joints) than the alternative (17 joints). The remainder of the approaches, on the other hand, work rather well on VideoPose3D. BlazePose adds unnecessary joint information that hinders model performance for [77], [109], [114], [115], whereas Deb et al. successfully ignores those via dynamic attention module.

B. Implementation Details

For models originally developed for assessing rehabilitation exercises (i.e. [87] & [77]), we closely followed the proposed architecture as described by the authors in the respective papers. On the other hand, for human activity based models ([109], [112], [113], [114], [115], and [117]), we replaced the last softmax layer with a fully-connected layer with linear activation in order to adapt them into the regression problem setting. We use the skeleton data as provided by the KIMORE

and UI-PRMD to construct the skeleton graph for graph-based approaches ([87], [109], [112], [113]). Similarly, for CNN-LSTM-based models ([114], [115], [117]), we also use skeleton data as input and transform these sequences as proposed by the respective authors. Regarding hyper-parameters, the batch size, the learning rate, and the dropout rate are selected from {16, 32, 64}, {2, 3, 4, 5, 6, 7}, and $\{10^{-2}, 10^{-3}, 10^{-4}\}$ via grid search [122]. All the models are trained with the Adam optimizer for a maximum of 100 epochs. The learning is terminated if the validation loss does not decrease in consecutive 30 epochs. All the models are initialized randomly and the results are reported by averaging over 10 runs.

For feature engineering-based methods, the features are selected based on [116] and [102]. We used Neural Network (NN) with the Adam optimizer for 500 epochs with a learning rate of 0.0001, two hidden layers with 12 and 24 neurons, respectively and Relu activation function. We also used Random Forest (RF) with two maximum depths, K-Nearest Neighbours (KNN) with 5 neighbors, and Support Vector Machine (SVM) with a radial basis function (RBF) kernel. We implement the machine learning algorithms (RF, KNN and SVM) in *sklearn*.¹ The NN models are implemented using Tensorflow.²

Discussion on RQ5

This section attempts to provide a quantitative analysis of some of the state-of-the-art works. Our analysis addresses the fifth research question regarding hand-crafted features and deep learning models. The results speak heavily in favor of the deep learning models. In each of the metrics, deep learning models performed considerably better than feature engineering-based methods. This thus proves conclusively that deep learning models should be the primary modality in future research.

¹<https://scikit-learn.org/>

²<https://www.tensorflow.org>

TABLE VIII

RESULTS OF EXERCISE 5 OF KIMORE WHILE COLLECTING JOINT/SKELETON DATA USING VARIOUS POSE ESTIMATION ALGORITHMS

Algorithm	Metric	Deep Learning					Feature Engineering-based							
		Deb et al. [87]	Liao et al. [77]	Yan et al. [109]	Du et al. [114]	Li et al. [115]	Capecci et al. [116]				Lee et al. [102]			
							K-NN	RF	SVM	NN	K-NN	RF	SVM	NN
BlazePose [119]	MAD	0.97	4.04	3.71	4.54	6.31	3.05	5.76	3.91	5.68	4.31	6.65	5.51	6.25
VideoPose3D [118]		1.85	2.55	3.08	3.54	4.66	3.33	4.87	3.54	4.12	5.35	7.42	6.00	6.77
Kinectv2 [9]		0.62	0.84	1.21	1.66	1.86	2.71	4.80	3.04	2.34	3.71	5.57	4.37	5.02
BlazePose [119]	RMSE	1.99	5.99	5.65	7.19	8.68	4.45	6.72	5.27	7.76	6.52	8.33	7.66	8.40
VideoPose3D [118]		3.82	3.91	4.94	5.20	6.01	5.05	6.32	5.21	5.81	7.40	9.60	8.65	9.77
Kinectv2 [9]		1.18	1.91	1.95	2.57	3.15	4.28	6.07	4.91	3.15	5.53	7.71	6.26	7.24
BlazePose [119]	MAPE	3.08	15.61	15.91	20.89	25.81	7.92	14.63	10.38	14.91	15.67	22.73	20.71	18.69
VideoPose3D [118]		6.81	8.10	10.79	11.96	14.75	9.05	13.70	10.33	10.42	19.61	27.90	24.46	23.21
Kinectv2 [9]		1.59	2.31	3.80	5.75	5.62	6.84	16.74	18.04	5.41	12.39	17.90	14.14	16.50

V. CHALLENGES IN AUTOMATIC STROKE REHABILITATION

Devising an automatic stroke rehabilitation system suffers from the unavailability of standard datasets, diversity in assessment methods, high implementation cost, and subject's safety concerns which hinder its successful implementation. In this section, we discuss these challenges in detail.

A. Inadequacy of Standard Data

Most stroke rehabilitation systems rely on data collected from local IoT devices such as Kinect V2, BTs Nirvana, and Vicon optical tracker, which are not publicly available. Few datasets provide action recognition data that are considerably different from rehabilitation exercises, particularly in terms of motion profiles. Again, the majority of datasets contain a limited number of samples from stroke patients, while others contain just healthy participants imitating various ailments [76]. Additionally, some datasets, such as UI-PRMD, are not scored by medical specialists.

B. Non-Uniformity In Assessment

The assessment mainly examines the severity of a stroke and the patient's functional ability, which a therapist does in a conventional system. This is highly subjective and frequently varies from therapist to therapist. Even among therapists, there is no universal agreement on using quantitative measures such as the FMA, ARAT, or other custom scoring systems for assessment purposes. The automated system also deals with the same challenges, and therefore a universal method of assessment is a must.

C. Increasing Cost of Rehabilitation

Our discussion, so far, indicates that automated systems are, as of now, much more expensive than conventional systems as the need for a skilled therapist still exists in both systems for purposes such as rehabilitation program design and final assessments. Automated systems provide superior performance and acceptability while reducing reliance on expert therapists, but at a higher expense. This increase in cost

can vary depending on the system. For example, the authors in [47] reported that VR-based systems can cost on average, £376 more than conventional therapy. Robotics-based systems can cost significantly more. According to [34], such a system could cost a patient between £666 and £1602 more than usual care, depending on the intensity of the program. Nonetheless, a trade-off exists between the expense of deployment and the efficacy of rehabilitation methods. However, as technology improves and costs decrease, automatic stroke rehabilitation systems may eventually supplant traditional rehabilitation.

D. Safety Concerns

Safety concerns in such automated systems is a big concern since it operates alongside patients in special care. For instance, one critical feature is automatically detecting when the patient or user of the system is experiencing discomfort [123]. Ergonomics of the devices used in the system is another area of concern, with many of the works, such as [45], [46], [59], and [66] reporting on the users' affinity with the system and ease of use. The system's stability is also a significant concern, given that the peripherals involved in the system are in very close proximity with the users. Although this is solely in the case of robotics aided systems, other safety concerns mentioned above apply to immersive VR based systems as well. Maintaining all of these safety features is challenging since they add complexity to the system, making mass deployment of such systems difficult.

Irrespective of that, getting over these hurdles and developing a safe-to-use rehabilitation system is crucial. Especially since these systems promote at-home rehabilitation where the patient is minimally supervised. Traditional inpatient rehabilitation therapy, despite its drawbacks, poses very few risk factors. Since the patients are constantly monitored, even if device malfunctions occur, they are immediately resolved by experts. Any sort of safety concerns can arise only at the time of human negligence and extenuating circumstances. Also, since the patient is observed and guided actively, negligence on the patient's part or misuse of assistive devices is minimal. On the contrary, at-home rehabilitation systems pose significant risks if not deployed correctly. Adding to the numerous safety concerns mentioned above, negligence and device misuse by

unsupervised patients can often create an unsafe environment and lead to sub-optimal results.

E. Biological Implications

Although this paper focuses primarily on the technological aspects of these automated systems, their biological implications should also be one of the main concerns. How such systems interact with the physic of a patient must be studied thoroughly before implementing such a system. Exoskeleton-based systems often conduct such studies before designing it [31], [39]. However, VR-based immersive systems rarely mention such a design consideration. This might be because such implications are less pertinent in their case. Nevertheless, studying these implications could provide insights into a more effective system overall.

Research shows that exoskeleton-based systems can reduce muscle volume [124] and muscle force [125]. According to these studies, exoskeletal assistance helps to activate muscles better and achieve more efficient movements. In order to get the optimal results, though, an early and intensive program should be considered [126]. The severity of the stroke affected also plays a part in rehabilitation [127]. One concern is that over reliance on such assistive exoskeletons could create problems with natural activation in the long run. In contrast to robot-assisted systems, VR-based systems affect the users on a psychological level as they engage the users to a great extent. This has great implications as it can support behavioral and habitual change [128], and increase their drive to complete the exercises successfully. All of the works reporting on VR-based systems demonstrate high user engagement and, in many cases, improved performance in daily tasks.

F. Adaptation of Automated Rehabilitation Systems

As reported in [45], [46], [58], and [59], users show a strong affinity for automated systems, despite some concerns about higher costs. Some works report on the needs and perspectives of the therapists involved in the deployment of these systems. For instance, authors in [129] discussed the key aspects for both improving patient movements and their progress assessment. In addition, many studies have acquired favor among physiotherapists as well [59] and [130]. Research is still ongoing, though, to find out which modality (robotics or VR) performs better, or which is more preferred.

Irrespective of that, there are also limitations to these automated rehabilitation systems. One work has reported on the poor generalization of results seen across such systems [131], which might not mirror the system efficacy. Skepticism about the rehabilitation system itself is also reported in the literature [132]. There also remains confusion and a lack of available information on how to effectively use the new technology [133]. Commercialization has also been cited as a hindrance [132], as this can lead to biased results. Finally, as of now, automated rehabilitation systems are quite costly. Recent research, however, predicts that the shift toward automating the rehabilitation system will reduce costs in the future [127], [134], and allow better outreach. Thus, despite these limitations, the world is moving in favor of such systems,

with the goal of providing affordable and available rehabilitation to all.

G. Discussion on RQ6

Our final research question inquired about the possibility to overcome the hurdles associated with automated systems. We have already discussed some of the works tailored to addressing these challenges. With constant breakthroughs and efforts in the field, it is safe to predict that these obstacles will be overcome in the near future.

VI. TAKEAWAYS

In this section, we present our findings according to the research questions posed in Section I.

RQ1: Our first research question challenged the scope and efficacy of automated rehabilitation systems. The literature review in Section II clearly shows that automated systems often exhibit better results than the conventional. Despite, the scopes of such systems are still in doubt due to the costs involved. Perhaps an automated system shared by a community would be the answer to this problem.

RQ2: Our subsequent research question raised concerns about the clinical validation. The standard approach towards this is to take a population and divide them into two groups, control and experimental. They are given therapy and the results are evaluated in different stages, using different metrics. From Table II and III, it is evident that most of the developed systems have either already completed clinical trials or awaiting one.

RQ3: Regarding the question of human involvement in automated systems, we have noticed that most of the works involve therapists in one way or another. Even though we are discussing automated systems, the exercise plan is always laid out by skilled professionals. Human supervisors are used in many of the works to assist and dictate the exercises, and also annotate datasets. **RQ4:** As to the answer to the following research question about the efficacy of such assessments, many of the works feature high agreement levels with the therapists' assessments [91], [92], [93], as well as high accuracy [94], [96], [97] and F1 scores [88], [91] in exercise correctness classification. This removes any doubts regarding the efficacy of such systems.

RQ5: Our fifth study topic posed an age-old question: is handmade feature-based learning better than deep learning methods? We show conclusive quantitative proof that when compared side by side on two standard datasets such as KIMORE (Table. VI) and UI-PRMD (Table. VII), deep learning algorithms are superior compared to the traditional hand-crafted. Especially in KIMORE, deep models conclusively outperform the hand-crafted methods in every exercise. In UI-PRMD dataset however, feature engineering based learning performs better than all the deep learning methods except for the work of Deb et al. [87]. The reasons for which has been discussed in detail in section IV.

RQ6: Our final query was whether we can move past the existing challenges. We have already discussed some of the

approaches to solving the associated problems. A collective approach can solve the inadequacy of data, which solely depends on the enthusiasm of the researchers involved. A similar statement could be made for the non-uniformity issue. Even the issue of increasing cost can be dealt with soon enough. However, dealing with the safety concerns will be a daunting task.

VII. CONCLUSION

The scope and promise of automated stroke rehabilitation systems with minimum human intervention are presented in our study. We have discussed cutting-edge methods for post-stroke rehabilitation, as well as data-driven exercise assessments. Our review of the literature suggests that automated rehabilitation systems have a lot of promise and can outperform traditional therapy. It has shown high levels of agreement with therapists' ratings in a number of studies. On two standard datasets, KIMORE and UI-PRMD, we offered a quantitative study of several state-of-the-art evaluation methods and a side-by-side comparison of deep learning and feature engineering-based learning approaches. The findings show that the deep models that incorporate spatio-temporal skeletal data and a dynamic attention module is extremely useful in rehabilitation assessment. We believe these automated technologies have the potential to make rehabilitation treatments more widely available. Future studies in this sector should undoubtedly focus on finding a cost-effective solution. Deep learning models should be investigated further, and an integrated strategy to gather real-world exercise data should be actively sought. Few shot [135] or zero shot [136] learning methods could be tried out in the absence of substantial datasets. These developments could eventually lead to a society where everyone has access to efficient rehabilitation services.

REFERENCES

- [1] C. M. Stinear, C. E. Lang, S. Zeiler, and W. D. Byblow, "Advances and challenges in stroke rehabilitation," *Lancet Neurol.*, vol. 19, no. 4, pp. 348–360, 2020.
- [2] M. Planton et al., "Neuropsychological outcome after a first symptomatic ischaemic stroke with 'good recovery,'" *Eur. J. Neurol.*, vol. 19, no. 2, pp. 212–219, 2012.
- [3] H. M. Dewey, L. J. Sherry, and J. M. Collier, "Stroke rehabilitation 2007: What should it be?" *Int. J. Stroke*, vol. 2, no. 3, pp. 191–200, Aug. 2007.
- [4] M. Gittler and A. M. Davis, "Guidelines for adult stroke rehabilitation and recovery," *JAMA*, vol. 319, no. 8, pp. 820–821, 2018.
- [5] J. P. Bettger, C. Liu, D. B. C. Gandhi, P. N. Sylaja, N. Jayaram, and J. D. Pandian, "Emerging areas of stroke rehabilitation research in low- and middle-income countries: A scoping review," *Stroke*, vol. 50, no. 11, pp. 3307–3313, Nov. 2019.
- [6] G. A. Roth et al., "Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: A systematic analysis for the global burden of disease study 2017," *Lancet*, vol. 392, no. 10159, pp. 1736–1788, 2018.
- [7] F. Gimigliano and S. Negrini, "The world health organization 'rehabilitation 2030: A call for action,'" *Eur. J. Phys. Rehabil. Med.*, vol. 53, no. 2, pp. 155–168, 2017.
- [8] B. H. Johnson, M. M. Bonafede, and C. Watson, "Short- and longer-term health-care resource utilization and costs associated with acute ischemic stroke," *Clinicoecon Outcomes Res.*, vol. 8, pp. 53–61, Feb. 2016.
- [9] J. Shotton et al., "Efficient human pose estimation from single depth images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 12, pp. 2821–2840, Dec. 2013.
- [10] T. Hussain, H. F. Maqbool, N. Iqbal, M. Khan, Salman, and A. A. Dehghani-Sanij, "Computational model for the recognition of lower limb movement using wearable gyroscope sensor," *Int. J. Sensor Netw.*, vol. 30, no. 1, pp. 35–45, 2019.
- [11] B. Pogorelec, Z. Bosnić, and M. Gams, "Automatic recognition of gait-related health problems in the elderly using machine learning," *Multimedia Tools Appl.*, vol. 58, no. 2, pp. 333–354, May 2012.
- [12] S. Das et al., "Quantitative measurement of motor symptoms in Parkinson's disease: A study with full-body motion capture data," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 6789–6792.
- [13] A. A. Frolov, I. B. Kozlovskaya, E. V. Biryukova, and P. D. Bobrov, "Use of robotic devices in post-stroke rehabilitation," *Neurosci. Behav. Physiol.*, vol. 48, no. 9, pp. 1053–1066, Nov. 2018.
- [14] N. Jarrasé et al., "Robotic exoskeletons: A perspective for the rehabilitation of arm coordination in stroke patients," *Frontiers Hum. Neurosci.*, vol. 8, p. 947, Dec. 2014.
- [15] V. Klamroth-Marganska, *Stroke Rehabilitation: Therapy Robots Assistive Devices*. Cham, Switzerland: Springer, 2018, pp. 579–587.
- [16] V. Klamroth-Marganska et al., "Three-dimensional, task-specific robot therapy of the arm after stroke: A multicentre, parallel-group randomised trial," *Lancet Neurol.*, vol. 13, no. 2, pp. 159–166, 2014.
- [17] H. Schmidt, J. Krüger, and S. Hesse, "HapticWalker—Haptic foot device for gait rehabilitation," in *Human Haptic Perception: Basics and Applications*. Basel, Switzerland: Springer, 2008, pp. 501–511.
- [18] P. Tamayo-Serrano, S. Garbaya, and P. Blazevic, "Gamified in-home rehabilitation for stroke survivors: Analytical review," *Int. J. Serious Games*, vol. 5, no. 1, pp. 2384–8766, Mar. 2018.
- [19] A. Gaber, M. F. Taher, and M. Waheb, "A comparison of virtual rehabilitation techniques," in *Proc. World Congr. (EECSS)*, Barcelona, Spain, 2015, pp. 13–14.
- [20] D. Webster and O. Celik, "Systematic review of Kinect applications in elderly care and stroke rehabilitation," *J. Neuroeng. Rehabil.*, vol. 11, no. 108, pp. 1–24, Jul. 2014.
- [21] J. Langan, H. Subryan, I. Nwogu, and L. Cavuoto, "Reported use of technology in stroke rehabilitation by physical and occupational therapists," *Disab. Rehabil., Assistive Technol.*, vol. 13, no. 7, pp. 641–647, 2018.
- [22] A. Rashid and O. Hasan, "Wearable technologies for hand joints monitoring for rehabilitation: A survey," *Microelectron. J.*, vol. 88, pp. 173–183, Jun. 2019.
- [23] D. Moher, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *Ann. Internal Med.*, vol. 151, no. 4, p. 264, Aug. 2009.
- [24] Y. Sankai, "HAL: Hybrid assistive limb based on cybernics," in *Robotics Research*. Berlin, Germany: Springer, 2010, pp. 25–34.
- [25] C. Yokota et al., "Acute stroke rehabilitation for gait training with cyborg type robot hybrid assistive limb: A pilot study," *J. Neurological Sci.*, vol. 404, pp. 11–15, Sep. 2019.
- [26] I. Aprile et al., "Robotic rehabilitation: An opportunity to improve cognitive functions in subjects with stroke. An explorative study," *Frontiers Neurol.*, vol. 11, p. 1498, Nov. 2020.
- [27] D. Wambach, M. Lamar, R. Swenson, D. Penney, E. Kaplan, and D. Libon, "Digit span," in *Encyclopedia of Clinical Neuropsychology*. New York, NY, USA: Springer, 2011, pp. 844–849.
- [28] N. Demeyerere, M. J. Riddoch, E. D. Slavkova, W.-L. Bickerton, and G. W. Humphreys, "The Oxford cognitive screen (OCS): Validation of a stroke-specific short cognitive screening tool," *Psychol. Assessment*, vol. 27, no. 3, pp. 883–894, 2015.
- [29] G. Demeurisse, O. Demol, and E. Robaye, "Motor evaluation in vascular hemiplegia," *Eur. Neurol.*, vol. 19, no. 6, pp. 382–389, 1980.
- [30] F. I. Mahoney et al., "Functional evaluation: The Barthel index," *Maryland State Med. J.*, vol. 14, no. 2, pp. 61–65, 1965.
- [31] Y.-K. Ou, Y.-L. Wang, H.-C. Chang, and C.-C. Chen, "Design and development of a wearable exoskeleton system for stroke rehabilitation," in *Healthcare*, vol. 8, no. 1. Basel, Switzerland: MDPI, 2020, p. 18.
- [32] A. Pilla et al., "Robotic rehabilitation and multimodal instrumented assessment of post-stroke elbow motor functions—A randomized controlled trial protocol," *Frontiers Neurol.*, vol. 11, p. 1294, Oct. 2020.

- [33] I. Aprile et al., "Upper limb robotic rehabilitation after stroke: A multicenter, randomized clinical trial," *J. Neurologic Phys. Therapy*, vol. 44, no. 1, pp. 3–14, 2020.
- [34] J. Bernhardt and J. Mehrholz, "Robotic-assisted training after stroke: RATULS advances science," *Lancet*, vol. 394, no. 10192, pp. 6–8, Jul. 2019.
- [35] Q. Qian, X. Hu, Q. Lai, S. C. Ng, Y. Zheng, and W. Poon, "Early stroke rehabilitation of the upper limb assisted with an electromyography-driven neuromuscular electrical stimulation-robotic arm," *Frontiers Neurol.*, vol. 8, p. 447, Sep. 2017.
- [36] B. R. W. and M. Smith, "Interrater reliability of a modified Ashworth scale of muscle spasticity," *Phys. Therapy*, vol. 67, no. 2, pp. 7–206, 1987.
- [37] N. Yozbatiran, L. Der-Yeghiaian, and S. C. Cramer, "A standardized approach to performing the action research arm test," *Neurorehabilitation Neural Repair*, vol. 22, no. 1, pp. 78–90, Jan. 2008.
- [38] J. L. Sullivan et al., "Improving robotic stroke rehabilitation by incorporating neural intent detection: Preliminary results from a clinical trial," in *Proc. Int. Conf. Rehabil. Robot. (ICORR)*, Jul. 2017, pp. 122–127.
- [39] J. Baca, M. S. Ambati, P. Dasgupta, and M. Mukherjee, "A modular robotic system for assessment and exercise of human movement," in *Advances in Automation and Robotics Research in Latin America*. Cham, Switzerland: Springer, 2017, pp. 61–70.
- [40] Z. Lu, K.-Y. Tong, X. Zhang, S. Li, and P. Zhou, "Myoelectric pattern recognition for controlling a robotic hand: A feasibility study in stroke," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 2, pp. 365–372, Feb. 2019.
- [41] K. Bustamante Valles, S. Montes, M. D. J. Madrigal, A. Burciaga, M. E. Martínez, and M. J. Johnson, "Technology-assisted stroke rehabilitation in Mexico: A pilot randomized trial comparing traditional therapy to circuit training in a robot/technology-assisted therapy gym," *J. NeuroEngineering Rehabil.*, vol. 13, no. 1, pp. 1–15, Dec. 2016.
- [42] N. Hogan, H. I. Krebs, A. Sharon, and J. Charnnarong, "Interactive robotic therapist," Massachusetts Institute of Technology, Cambridge, MA, USA, U.S. Patent 5466, p. 466, 1995.
- [43] M. Gilliaux et al., "Using the robotic device REAplan as a valid, reliable, and sensitive tool to quantify upper limb impairments in stroke patients," *J. Rehabil. Med.*, vol. 46, no. 2, pp. 117–125, 2014.
- [44] H. I. Krebs, N. Hogan, M. L. Aisen, and B. T. Volpe, "Robot-aided neurorehabilitation," *IEEE Trans. Rehabil. Eng.*, vol. 6, no. 1, pp. 75–87, Mar. 1998.
- [45] S. Heins et al., "Robotic-assisted serious game for motor and cognitive post-stroke rehabilitation," in *Proc. IEEE 5th Int. Conf. Games Appl. Health (SeGAH)*, Apr. 2017, pp. 1–8.
- [46] K. J. Bower, J. Louie, Y. Landesrocha, P. Seedy, A. Gorelik, and J. Bernhardt, "Clinical feasibility of interactive motion-controlled games for stroke rehabilitation," *J. Neuroeng. Rehabil.*, vol. 12, no. 1, pp. 1–12, Dec. 2015.
- [47] K. Adie et al., "Does the use of Nintendo Wii Sports™ improve arm function? Trial of Wii™ in stroke: A randomized controlled trial and economics analysis," *Clin. Rehabil.*, vol. 31, no. 2, pp. 173–185, Feb. 2017.
- [48] O. Rosiak, K. Krajewski, M. Woszczak, and M. Jozefowicz-Korczyńska, "Evaluation of the effectiveness of a virtual reality-based exercise program for unilateral peripheral vestibular deficit," *J. Vestibular Res.*, vol. 28, nos. 5–6, pp. 409–415, Apr. 2019.
- [49] L. Yardley, E. Masson, C. Verschuur, N. Haacke, and L. Luxon, "Symptoms, anxiety and handicap in dizzy patients: Development of the vertigo symptom scale," *J. Psychosomatic Res.*, vol. 36, no. 8, pp. 731–741, Dec. 1992.
- [50] C. Schuster-Amft et al., "Effect of a four-week virtual reality-based training versus conventional therapy on upper limb motor function after stroke: A multicenter parallel group randomized trial," *PLoS ONE*, vol. 13, no. 10, Oct. 2018, Art. no. e0204455.
- [51] V. Mathiowetz, G. Volland, N. Kashman, and K. Weber, "Adult norms for the box and block test of manual dexterity," *Amer. J. Occupat. Therapy*, vol. 39, pp. 386–391, Jun. 1985.
- [52] S. Barreca et al., "Development of the chedoke arm and hand activity inventory: Theoretical constructs, item generation, and selection," *Top. Stroke Rehabil.*, vol. 11, no. 4, pp. 31–42, 2004.
- [53] P. W. Duncan, D. Wallace, S. M. Lai, D. Johnson, S. Embretson, and L. J. Laster, "The stroke impact scale version 2.0: Evaluation of reliability, validity, and sensitivity to change," *Stroke*, vol. 30, no. 10, pp. 2131–2140, Oct. 1999.
- [54] T.-H. Ho et al., "Impact of virtual reality-based rehabilitation on functional outcomes in patients with acute stroke: A retrospective case-matched study," *J. Neurol.*, vol. 266, no. 3, pp. 589–597, Mar. 2019.
- [55] S. Y. Joo, Y. S. Cho, S. Y. Lee, H. Seok, and C. H. Seo, "Effects of virtual reality-based rehabilitation on burned hands: A prospective, randomized, single-blind study," *J. Clin. Med.*, vol. 9, no. 3, p. 731, Mar. 2020.
- [56] R. H. Jebsen, N. Taylor, R. B. Trieschmann, M. J. Trotter, and L. A. Howard, "An objective and standardized test of hand function," *Arch. Phys. Med. Rehabil.*, vol. 50, no. 6, pp. 311–319, Jun. 1969.
- [57] K. C. Chung, M. S. Pillsbury, M. R. Walters, and R. A. Hayward, "Reliability and validity testing of the Michigan hand outcomes questionnaire," *J. Hand Surgery*, vol. 23, no. 4, pp. 575–587, Jul. 1998.
- [58] A. Warland et al., "The feasibility, acceptability and preliminary efficacy of a low-cost, virtual-reality based, upper-limb stroke rehabilitation device: A mixed methods study," *Disability Rehabil.*, vol. 41, no. 18, pp. 2119–2134, Aug. 2019.
- [59] A. Elor, S. Kurniawan, and M. Teodorescu, "Towards an immersive virtual reality game for smarter post-stroke rehabilitation," in *Proc. IEEE Int. Conf. Smart Comput. (SMARTCOMP)*, Jun. 2018, pp. 219–225.
- [60] E. Taub et al., "Constraint-induced movement therapy: A new family of techniques with broad application to physical rehabilitation—A clinical review," *J. Rehabil. Res.*, vol. 36, no. 3, pp. 237–251, Jul. 1999.
- [61] S. J. Page, P. Levine, S. Sisto, Q. Bond, and M. V. Johnston, "Stroke patients' and therapists' opinions of constraint-induced movement therapy," *Clin. Rehabil.*, vol. 16, no. 1, pp. 55–60, Feb. 2002.
- [62] R. De Luca et al., "Effects of virtual reality-based training with BTsnirvana on functional recovery in stroke patients: Preliminary considerations," *Int. J. Neurosci.*, vol. 128, no. 9, pp. 791–796, Sep. 2018.
- [63] T. Prasertsakul, P. Kaimuk, W. Chinjenpradit, W. Limroongreungrat, and W. Charoensuk, "The effect of virtual reality-based balance training on motor learning and postural control in healthy adults: A randomized preliminary study," *Biomed. Eng. OnLine*, vol. 17, no. 1, pp. 1–17, Dec. 2018.
- [64] A. Vourvopoulos, O. Marin-Pardo, M. Neureither, D. Saldana, E. Jahng, and S.-L. Liew, "Multimodal head-mounted virtual-reality brain-computer interface for stroke rehabilitation," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2019, pp. 165–179.
- [65] R. Spicer, J. Anglin, D. M. Krum, and S.-L. Liew, "REINVENT: A low-cost, virtual reality brain-computer interface for severe stroke upper limb motor recovery," in *Proc. IEEE Virtual Reality (VR)*, Mar. 2017, pp. 385–386.
- [66] O. Marin-Pardo, A. Vourvopoulos, M. Neureither, D. Saldana, E. Jahng, and S.-L. Liew, "Electromyography as a suitable input for virtual reality-based biofeedback in stroke rehabilitation," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2019, pp. 274–281.
- [67] A. Miron, N. Sadawi, W. Ismail, H. Hussain, and C. Grosan, "IntelliRehabDS (IRDS)—A dataset of physical rehabilitation movements," *Data*, vol. 6, no. 5, p. 46, Apr. 2021.
- [68] F. Sardari, A. Paiement, S. Hannuna, and M. Mirmehdi, "VI-Net—View-invariant quality of human movement assessment," *Sensors*, vol. 20, no. 18, p. 5258, Sep. 2020.
- [69] A. Wijekoon, N. Wiratunga, and K. Cooper, "MEx: Multi-modal exercises dataset for human activity recognition," 2019, *arXiv:1908.08992*.
- [70] M. Capecci et al., "The KIMORE dataset: Kinematic assessment of Movement and clinical scores for remote monitoring of physical Rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 7, pp. 1436–1448, Jul. 2019.
- [71] M. Capecci et al., "An instrumental approach for monitoring physical exercises in a visual markerless scenario: A proof of concept," *J. Biomechanics*, vol. 69, pp. 70–80, Mar. 2018.
- [72] A. Wijekoon, N. Wiratunga, and K. Cooper, "Heterogeneous multimodal sensor fusion with hybrid attention for exercise recognition," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–8.
- [73] A. Wijekoon, N. Wiratunga, K. Cooper, and K. Bach, "Learning to recognise exercises in the self-management of low back pain," in *Proc. 33rd Int. Flairs Conf.*, 2020, pp. 1–6.
- [74] A. Wijekoon and N. Wiratunga, "Personalised meta-learning for human activity recognition with few-data," in *Proc. Int. Conf. Innov. Techn. Appl. Artif. Intell.* Cham, Switzerland: Springer, 2020, pp. 79–93.
- [75] Q. Guo and S. Khan, "Exercise-specific feature extraction approach for assessing physical rehabilitation," in *Proc. 4th Workshop AI Aging, Rehabil., Intell. Assist. Living, Int. Joint Conf. Artif. Intell. (IJCAI)*, 2021.
- [76] A. Vakanski, H.-P. Jun, D. Paul, and R. Baker, "A data set of human body movements for physical rehabilitation exercises," *Data*, vol. 3, no. 1, p. 2, Jan. 2018.

- [77] Y. Liao, A. Vakanski, and M. Xian, "A deep learning framework for assessing physical rehabilitation exercises," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 468–477, Feb. 2020.
- [78] L. Li and A. Vakanski, "Generative adversarial networks for generation and classification of physical rehabilitation movement episodes," *Int. J. Mach. Learn. Comput.*, vol. 8, no. 5, pp. 428–436, 2018.
- [79] B. X. B. Yu, Y. Liu, K. C. C. Chan, Q. Yang, and X. Wang, "Skeleton-based human action evaluation using graph convolutional network for monitoring Alzheimer's progression," *Pattern Recognit.*, vol. 119, Nov. 2021, Art. no. 108095.
- [80] F. A. N. Rashid, N. S. Suriani, M. N. Mohd, M. R. Tomari, W. N. W. Zakaria, and A. Nazari, "Deep convolutional network approach in spike train analysis of physiotherapy movements," in *Advances in Electronics Engineering*. Singapore: Springer, 2020, pp. 159–170.
- [81] X. Yu, "Practical algorithms for vision-based human activity recognition and human action evaluation," Ph.D. dissertation, Dept. Comput., Hong Kong Polytech. Univ., Hong Kong, 2020. [Online]. Available: <https://theses.lib.polyu.edu.hk/handle/200/11050>
- [82] W. Zhang, C. Su, and C. He, "Rehabilitation exercise recognition and evaluation based on smart sensors with deep learning framework," *IEEE Access*, vol. 8, pp. 77561–77571, 2020.
- [83] J. Antunes, A. Bernardino, A. Smailagic, and D. P. Siewiorek, "AHA-3D: A labelled dataset for senior fitness exercise recognition and segmentation from 3D skeletal data," in *Proc. BMVC*, 2018, p. 332.
- [84] E. Dolatabadi et al., "The Toronto rehab stroke pose dataset to detect compensation during stroke rehabilitation therapy," in *Proc. 11th EAI Int. Conf. Pervasive Comput. Technol. Healthcare*, 2017, pp. 375–381.
- [85] X. Z. Ying, M. Lukasik, M. H. Li, E. Dolatabadi, R. H. Wang, and B. Taati, "Automatic detection of compensation during robotic stroke rehabilitation therapy," *IEEE J. Transl. Eng. Health Med.*, vol. 6, pp. 1–7, 2017.
- [86] S. Rich U. Uy and P. Angela Abu, "Analysis of detecting compensation for robotic stroke rehabilitation therapy using imbalanced learning and outlier detection," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIC)*, Feb. 2020, pp. 432–437.
- [87] S. Deb, M. F. Islam, S. Rahman, and S. Rahman, "Graph convolutional networks for assessment of physical rehabilitation exercises," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 410–419, 2022.
- [88] S. Cai et al., "Detecting compensatory movements of stroke survivors using pressure distribution data and machine learning algorithms," *J. Neuroeng. Rehabil.*, vol. 16, no. 1, pp. 1–11, Dec. 2019.
- [89] H.-T. Jung, J. Park, J. Jeong, T. Ryu, Y. Kim, and S. I. Lee, "A wearable monitoring system for at-home stroke rehabilitation exercises: A preliminary study," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Mar. 2018, pp. 13–16.
- [90] E. Frank, Y. Wang, S. Inglis, G. Holmes, and I. H. Witten, "Using model trees for classification," *Mach. Learn.*, vol. 32, no. 1, pp. 63–76, 1998.
- [91] M. H. Lee, "An intelligent decision support system for stroke rehabilitation assessment," in *Proc. 21st Int. ACM SIGACCESS Conf. Comput. Accessibility*, 2019, pp. 694–696.
- [92] M. H. Lee, D. P. Siewiorek, A. Smailagic, A. Bernardino, and S. Bermúdez I Badia, "An exploratory study on techniques for quantitative assessment of stroke rehabilitation exercises," in *Proc. 28th ACM Conf. User Model., Adaptation Personalization*, Jul. 2020, pp. 303–307.
- [93] M. H. Lee, D. P. Siewiorek, A. Smailagic, A. Bernardino, and S. Bermúdez I Badia, "Co-design and evaluation of an intelligent decision support system for stroke rehabilitation assessment," in *Proc. ACM Hum.-Comput. Interact.*, Oct. 2020, vol. 4, no. CSCW2, pp. 1–27.
- [94] C. Li, L. Cheng, H. Yang, Y. Zou, and F. Huang, "An automatic rehabilitation assessment system for hand function based on leap motion and ensemble learning," *Cybern. Syst.*, vol. 52, no. 1, pp. 3–25, Jan. 2021.
- [95] A. Kaku, A. Parnandi, A. Venkatesan, N. Pandit, H. Schambra, and C. Fernandez-Granda, "Towards data-driven stroke rehabilitation via wearable sensors and deep learning," in *Proc. Mach. Learn. Healthcare Conf.*, 2020, pp. 143–171.
- [96] M. Panwar et al., "Rehab-Net: Deep learning framework for arm movement classification using wearable sensors for stroke rehabilitation," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 11, pp. 3026–3037, Nov. 2019.
- [97] Z.-A. Zhu, Y.-C. Lu, C.-H. You, and C.-K. Chiang, "Deep learning for sensor-based rehabilitation exercise recognition and evaluation," *Sensors*, vol. 19, no. 4, p. 887, Feb. 2019.
- [98] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, "Analysis of human behavior recognition algorithms based on acceleration data," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2013, pp. 1602–1607.
- [99] Z. Zhang, Q. Fang, and X. Gu, "Objective assessment of upper-limb mobility for poststroke rehabilitation," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 859–868, Apr. 2016.
- [100] A. Goñi, A. Illarramendi, and D. Antón, "Exercise recognition for Kinect-based telerehabilitation," *Methods Inf. Med.*, vol. 54, no. 2, pp. 145–155, 2015.
- [101] J. Wang, L. Yu, J. Wang, L. Guo, X. Gu, and Q. Fang, "Automated Fugl-Meyer assessment using SVR model," in *Proc. IEEE Int. Symp. Bioelectronics Bioinf.*, Apr. 2014, pp. 1–4.
- [102] M. H. Lee, D. P. Siewiorek, A. Smailagic, A. Bernardino, and S. B. I. Badia, "Learning to assess the quality of stroke rehabilitation exercises," in *Proc. 24th Int. Conf. Intell. User Interface*, Mar. 2019, pp. 218–228.
- [103] A. Yurtman and B. Barshan, "Automated evaluation of physical therapy exercises using multi-template dynamic time warping on wearable sensor signals," *Comput. Methods Programs Biomed.*, vol. 117, no. 2, pp. 189–207, Nov. 2014.
- [104] M. Müller, "Dynamic time warping," in *Information Retrieval for Music and Motion*. Berlin, Germany: Springer, 2007, pp. 69–84.
- [105] S.-W. Kim, S. Park, and W. W. Chu, "An index-based approach for similarity search supporting time warping in large sequence databases," in *Proc. 17th Int. Conf. Data Eng.*, 2001, pp. 607–614.
- [106] C.-J. Su, C.-Y. Chiang, and J.-Y. Huang, "Kinect-enabled home-based rehabilitation system using dynamic time warping and fuzzy logic," *Appl. Soft Comput.*, vol. 22, no. Nov. 2010, pp. 652–666, Sep. 2014.
- [107] S. Schetz-Sobrino, D. N. Monekoso, P. Remagnino, D. Vallejo, and C. Glez-Morcillo, "Automatic recognition of physical exercises performed by stroke survivors to improve remote rehabilitation," in *Proc. Int. Conf. Multimedia Anal. Pattern Recognit. (MAPR)*, May 2019, pp. 1–6.
- [108] S. Schetz-Sobrino, D. Vallejo, D. N. Monekoso, C. Glez-Morcillo, and P. Remagnino, "A distributed gamified system based on automatic assessment of physical exercises to promote remote physical rehabilitation," *IEEE Access*, vol. 8, pp. 91424–91434, 2020.
- [109] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition," in *Proc. AAAI*, vol. 32, no. 1, 2018, pp. 7444–7452.
- [110] Z. Liu, H. Zhang, Z. Chen, Z. Wang, and W. Ouyang, "Disentangling and unifying graph convolutions for skeleton-based action recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 143–152.
- [111] F. Ye, S. Pu, Q. Zhong, C. Li, D. Xie, and H. Tang, "Dynamic GCN: Context-enriched topology learning for skeleton-based action recognition," in *Proc. 28th ACM Int. Conf. Multimedia*, Oct. 2020, pp. 55–63.
- [112] Y. F. Song, Z. Zhang, C. Shan, and L. Wang, "Richly activated graph convolutional network for robust skeleton-based action recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 5, pp. 1915–1925, May 2021.
- [113] P. Zhang, C. Lan, W. Zeng, J. Xing, J. Xue, and N. Zheng, "Semantics-guided neural networks for efficient skeleton-based human action recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1112–1121.
- [114] Y. Du, W. Wang, and L. Wang, "Hierarchical recurrent neural network for skeleton based action recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1110–1118.
- [115] C. Li, Q. Zhong, D. Xie, and S. Pu, "Co-occurrence feature learning from skeleton data for action recognition and detection with hierarchical aggregation," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Jul. 2018, pp. 786–792.
- [116] M. Capecci et al., "The KIMORE dataset: Kinematic assessment of movement and clinical scores for remote monitoring of physical rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 7, pp. 1436–1448, Jul. 2019.
- [117] A. Shahroudy, J. Liu, T.-T. Ng, and G. Wang, "NTU RGB+D: A large scale dataset for 3D human activity analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1010–1019.
- [118] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3D human pose estimation in video with temporal convolutions and semi-supervised training," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 7753–7762.

- [119] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, and M. Grundmann, "BlazePose: On-device real-time body pose tracking," in *Proc. CVPR Workshop Comput. Vis. Augmented Virtual Reality*, 2020, pp. 1–4.
- [120] T.-Y. Lin et al., "Microsoft COCO: Common objects in context," in *Proc. ECCV*, 2014, pp. 740–755.
- [121] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: Large scale datasets and predictive methods for 3D human sensing in natural environments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 7, pp. 1325–1339, Dec. 2013.
- [122] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [123] P. Werner, A. Al-Hamadi, K. Limbrecht-Ecklundt, S. Walter, and H. C. Traue, "Head movements and postures as pain behavior," *PLoS ONE*, vol. 13, no. 2, Feb. 2018, Art. no. e0192767.
- [124] O. N. Beck, L. K. Punith, R. W. Nuckols, and G. S. Sawicki, "Exoskeletons improve locomotion economy by reducing active muscle volume," *Exercise Sport Sci. Rev.*, vol. 47, no. 4, pp. 237–245, 2019.
- [125] B. J. McLain, D. Lee, S. C. Mulrine, and A. J. Young, "Effect of assistance using a bilateral robotic knee exoskeleton on tibiofemoral force using a neuromuscular model," *Ann. Biomed. Eng.*, vol. 50, no. 6, pp. 716–727, Jun. 2022.
- [126] J. Mehrholz, A. Hädrich, T. Platz, J. Kugler, and M. Pohl, "Electromechanical and robot-assisted arm training for improving generic activities of daily living, arm function, and arm muscle strength after stroke," *Cochrane Database Systematic Rev.*, no. 6, Jun. 2012.
- [127] S. Masiero et al., "The value of robotic systems in stroke rehabilitation," *Exp. Rev. Med. Devices*, vol. 11, no. 2, pp. 187–198, Mar. 2014.
- [128] Y. Meng and S. S. Wong, "Implications of virtual reality technology in nutrition intervention (811.28)," *FASEB J.*, vol. 28, no. S1, pp. 28–811, Apr. 2014.
- [129] A. W. K. Lam, D. Varona-Marin, Y. Li, M. Fergenbaum, and D. Kulić, "Automated rehabilitation system: Movement measurement and feedback for patients and physiotherapists in the rehabilitation clinic," *Human Computer Interact.*, vol. 31, nos. 3–4, pp. 294–334, Jul. 2016.
- [130] K. Chua et al., "An exploratory clinical study on an automated, speed-sensing treadmill prototype with partial body weight support for hemiparetic gait rehabilitation in subacute and chronic stroke patients," *Frontiers Neurol.*, vol. 11, p. 747, Jul. 2020.
- [131] G. Morone et al., "Who may have durable benefit from robotic gait training? A 2-year follow-up randomized controlled trial in patients with subacute stroke," *Stroke*, vol. 43, no. 4, pp. 1140–1142, 2012.
- [132] M. Iosa, G. Morone, A. Cherubini, and S. Paolucci, "The three laws of neurorobotics: A review on what neurorehabilitation robots should do for patients and clinicians," *J. Med. Biol. Eng.*, vol. 36, no. 1, pp. 1–11, Feb. 2016.
- [133] B. Sheng, Y. Zhang, W. Meng, C. Deng, and S. Xie, "Bilateral robots for upper-limb stroke rehabilitation: State of the art and future prospects," *Med. Eng. Phys.*, vol. 38, no. 7, pp. 587–606, Jul. 2016.
- [134] E. Datteri, "Predicting the long-term effects of human–robot interaction: A reflection on responsibility in medical robotics," *Sci. Eng. Ethics*, vol. 19, no. 1, pp. 139–160, Mar. 2013.
- [135] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few-shot learning," *ACM Comput. Surv.*, vol. 53, no. 3, pp. 1–34, 2020.
- [136] W. Wang, V. W. Zheng, H. Yu, and C. Miao, "A survey of zero-shot learning: Settings, methods, and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–37, 2019.